

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7: C12N 15/52, 15/54, 15/62, 9/10, C12P 17/18, 19/32, C07D 498/18 // (C07D 498/18, 311:00, 273:00, 211:00) (11) International Publication Number:

WO 00/20601

(43) International Publication Date:

13 April 2000 (13.04.00)

(21) International Application Number:

PCT US99 22886

A2

(22) International Filing Date:

1 October 1999 (01.10.99)

(74) Agents: FAVORITO, Carolyn et al.; Morrison & Foerster LLP, 2000 Pennsylvania Avenue, N.W., Washington, DC 20006–1888 (US).

(30) Priority Data:

60/102,748 2 October 1998 (02.10.98) US 60/123,810 11 March 1999 (11.03.99) US 60'139,650 17 June 1999 (17.06.99) US (81) Designated States: AL, AM, AU, BA, BB, BG, BR, CA, CN, CR, CU, CZ, DM, EE, GD, GE, HR, HU, IL, IS, JP, KG, KP, KR, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, US, UZ, VN, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant 'for all designated States except US: KOSAN BIOSCIENCES, INC. [US/US]: 3832 Bay Center Drive, Hayward, CA 94545 (US).

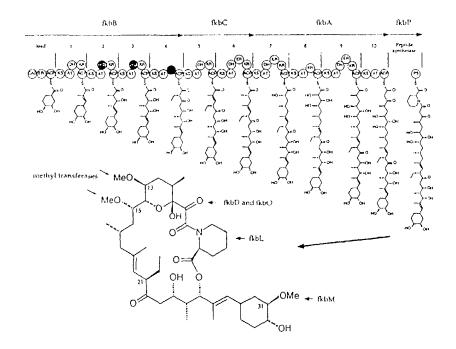
(72) Inventors; and

(75) Inventors/Applicants (for US only): REEVES, Christopher [US/US]; 4 East Altarinda Drive, Orinda, CA 94563 (US), CHU, Daniel [US/US]; 3767 Benton Street, Santa Clara, CA 95051 (US), KHOSLA, Chaitan [IN US]; 740 Para Avenue, Palo Alto, CA 94306 (US), SANTI, Daniel [US/US]; 211 Belgrave Avenue, San Francisco, CA 94117 (US), WU, Kai [CN/US]; 900 Constitution Drive, Foster City, CA 94404 (US).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA CONSTRUCTS THEREFOR



(57) Abstract

Host cells comprising recombinant vectors encoding the FK-520 polyketide synthase and FK-520 modification enzymes can be used to produce the FK-520 polyketide. Recombinant DNA constructs comprising one or more FK-520 polyketide synthase domains, modules, open reading frames, and variants thereof can be used to produce recombinant polyketide synthases and a variety of different polyketides with application as pharmaceutical and veterinary products.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

ΑL	Albama	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Luthuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Scnegal
\mathbf{AU}	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
вв	Barbados	GH	Ghana	MG	Madagascar	ТJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
$\mathbf{B}F$	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	Τ̃Γ	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	U.	Israel	MR	Mauritania	UG	Uganda
BY	Belanis	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JР	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Nuzakstan	RO	Romania		
CZ	Czech Republic	1.C	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA CONSTRUCTS THEREFOR

Field of the Invention

5

10

15

20

25

30

35

The present invention relates to polyketides and the polyketide synthase (PKS) enzymes that produce them. The invention also relates generally to genes encoding PKS enzymes and to recombinant host cells containing such genes and in which expression of such genes leads to the production of polyketides. The present invention also relates to compounds useful as medicaments having immunosuppressive and/or neurotrophic activity. Thus, the invention relates to the fiel of chemistry, molecular biology, and agricultural, medical, and veterinary technology.

Background of the Invention

Polyketides are a class of compounds synthesized from 2-carbon units through a series of condensations and subsequent modifications. Polyketides occur in many types of organisms, including fungi and mycelial bacteria, in particular, the actinomycetes. Polyketides are biologically active molecules with a wide variety of structures, and the class encompasses numerous compounds with diverse activities. Tetracycline, erythromycin, epothilone, FK-506, FK-520, narbomycin, picromycin, rapamycin, spinocyn, and tylosin are examples of polyketides. Given the difficulty in producing polyketide compounds by traditional chemical methodology, and the typically low production of polyketides in wild-type cells, there has been considerable interest in finding improved or alternate means to produce polyketide compounds.

This interest has resulted in the cloning, analysis, and manipulation by recombinant DNA technology of genes that encode PKS enzymes. The resulting technology allows one to manipulate a known PKS gene cluster either to produce the polyketide synthesized by that PKS at higher levels than occur in nature or in hosts that otherwise do not produce the polyketide. The technology also allows one to produce molecules that are structurally related to, but distinct from, the polyketides produced from known PKS gene clusters. See, e.g., PCT publication Nos. WO 93/13663; 95/08548; 96/40968; 97/02358; 98/27203; and 98/49315; United States Patent Nos. 4,874,748; 5,063,155; 5,098,837; 5,149,639; 5,672,491; 5,712,146; 5,830,750; and 5,843,718; and Fu et al., 1994, Biochemistry 33: 9321-9326; McDaniel et al., 1993, Science 262: 1546-1550; and Rohr, 1995, Angew. Chem. Int. Ed. Engl. 34(8): 881-888, each of which is incorporated herein by reference.

Polyketides are synthesized in nature by PKS enzymes. These enzymes, which are complexes of multiple large proteins, are similar to the synthases that catalyze condensation of 2-carbon units in the biosynthesis of fatty acids. PKSs catalyze the biosynthesis of polyketides through repeated, decarboxylative Claisen condensations between acylthioester building blocks. The building blocks used to form complex polyketides are typically acylthioesters, such as acetyl, butyryl, propionyl, malonyl, hydroxymalonyl, methylmalonyl, and ethylmalonyl CoA. Other building blocks include amino acid like acylthioesters. PKS enzymes that incorporate such building blocks include an activity that functions as an amino acid ligase (an AMP ligase) or as a non-ribosomal patide synthetase (NRPS). Two major types of PKS enzymes are known; these differ in their composition and mode of synthesis of the polyketide synthesized. These two major types of PKS enzymes are commonly referred to as Type I or "modular" and Type II "iterative" PKS enzymes.

In the Type I or modular PKS enzyme group, a set of separate catalytic active sites (each active site is termed a "domain", and a set thereof is termed a "module") exists for each cycle of carbon chain elongation and modification in the polyketide synthesis pathway. The typical modular PKS is composed of several large polypeptides, which can be segregated from amino to carboxy termini into a loading module, multiple extender modules, and a releasing (or thioesterase) domain. The PKS enzyme known as 6-deoxyerythronolide B synthase (DEBS) is a Type I PKS. In DEBS, there is a loading module, six extender modules, and a thioesterase (TE) domain. The loading module, six extender modules, and TE of DEBS are present on three separate proteins (designated DEBS-1, DEBS-2, and DEBS-3, with two extender modules per protein). Each of the DEBS polypeptides is encoded by a separate open reading frame (ORF) or gene; these genes are known as *eryAI*, *eryAII*, and *eryAIII*. See Caffrey *et al.*, 1992, *FEBS Letters* 304: 205, and U.S. Patent No. 5,824,513, each of which is incorporated herein by reference.

Generally, the loading module is responsible for binding the first building block used to synthesize the polyketide and transferring it to the first extender module. The loading module of DEBS consists of an acyltransferase (AT) domain and an acyl carrier protein (ACP) domain. Another type of loading module utilizes an inactivated ketosynthase (KS) domain and AT and ACP domains. This inactivated KS is in some instances called KS^Q, where the superscript letter is the abbreviation for the amino acid, glutainine, that is present instead of the active site cysteine required for ketosynthase activity. In other PKS enzymes, including the FK-506 PKS, the loading module

incorporates an unusual starter unit and is composed of a CoA ligase like activity domain. In any event, the loading module recognizes a particular acyl-CoA (usually acetyl or propionyl but sometimes butyryl or other acyl-CoA) and transfers it as a thiol ester to the ACP of the loading module.

5

Đ

15

20

25

30

The AT on each of the extender modules recognizes a particular extender-CoA (malonyl or alpha-substituted malonyl, i.e., methylmalonyl, ethylmalonyl, and 2-hydroxymalonyl) and transfers it to the ACP of that extender module to form a thioester. Each extender module is responsible for accepting a compound from a prior module, binding a building block, attaching the building block to the compound from the prior module, optionally performing one or more additional range one, and transferring the resulting compound to the next module.

Each extender module of a modular PKS contains a KS. AT. ACP, and zero, one, two, or three domains that modify the beta-carbon of the growing polyketide chain. A typical (non-loading) minimal Type I PKS extender module is exemplified by extender module three of DEBS, which contains a KS domain, an AT domain, and an ACP domain. These three domains are sufficient to activate a 2-carbon extender unit and attach it to the growing polyketide molecule. The next extender module, in turn, is responsible for attaching the next building block and transferring the growing compound to the next extender module until synthesis is complete.

Once the PKS is primed with acyl- and malonyl-ACPs, the acyl group of the loading module is transferred to form a thiol ester (trans-esterification) at the KS of the first extender module; at this stage, extender module one possesses an acyl-KS and a malonyl (or substituted malonyl) ACP. The acyl group derived from the loading module is then covalently attached to the alpha-carbon of the malonyl group to form a carbon-carbon bond, driven by concomitant decarboxylation, and generating a new acyl-ACP that has a backbone two carbons longer than the loading building block (elongation or extension).

The polyketide chain, growing by two carbons each extender module, is sequentially passed as covalently bound thiol esters from extender module to extender module, in an assembly line-like process. The carbon chain produced by this process alone would possess a ketone at every other carbon atom. producing a polyketone, from which the name polyketide arises. Most commonly, however, additional enzymatic activities modify the beta keto group of each two carbon unit just after it has been added to the growing polyketide chain but before it is transferred to the next module.

Thus, in addition to the minimal module containing KS, AT, and ACP domains necessary to form the carbon-carbon bond, and as noted above, other domains that modify the beta-carbonyl moiety can be present. Thus, modules may contain a ketoreductase (KR) domain that reduces the keto group to an alcohol. Modules may also contain a KR domain plus a dehydratase (DH) domain that dehydrates the alcohol to a double bond. Modules may also contain a KR domain, a DH domain, and an enoylreductase (ER) domain that converts the double bond product to a saturated single bond using the beta carbon as a methylene function. An extender module can also contain other enzymatic activity, such as, for example, a methylase or dimethylase activity.

After traversing the final extender module, the polyketide encounters a releasing domain that cleaves the polyketide from the PKS and typically cyclizes the polyketide. For example, final synthesis of 6-dEB is regulated by a TE domain located at the end of extender module six. In the synthesis of 6-dEB, the TE domain catalyzes cyclization of the macrolide ring by formation of an ester linkage. In FK-506, FK-520, rapamycin, and similar polyketides, the TE activity is replaced by a RapP (for rapamycin) or RapP like activity that makes a linkage incorporating a pipecolate acid residue. The enzymatic activity that catalyzes this incorporation for the rapamycin enzyme is known as RapP, encoded by the *rapP* gene. The polyketide can be modified further by tailoring enzymes; these enzymes add carbohydrate groups or methyl groups, or make other modifications. i.e., oxidation or reduction, on the polyketide core molecule. For example, 6-dEB is hydroxylated at C-6 and C-12 and glycosylated at C-3 and C-5 in the synthesis of ervthromycin A.

In Type I PKS polypeptides, the order of catalytic domains is conserved. When all beta-keto processing domains are present in a module, the order of domains in that module from N-to-C-terminus is always KS, AT, DH, ER, KR, and ACP. Some or all of the beta-keto processing domains may be missing in particular modules, but the order of the domains present in a module remains the same. The order of domains within modules is believed to be important for proper folding of the PKS polypetides into an active complex. Importantly, there is considerable flexibility in PKS enzymes, which allows for the genetic engineering of novel catalytic complexes. The engineering of these enzymes is achieved by modifying, adding, or deleting domains, or replacing them with those taken from other Type I PKS enzymes. It is also achieved by deleting, replacing, or adding entire modules with those taken from other sources. A genetically engineered

PKS complex should of course have the ability to catalyze the synthesis of the product predicted from the genetic alterations made.

5

10

15

20

25

30

35

Alignments of the many available amino acid sequences for Type I PKS enzymes has approximately defined the boundaries of the various catalytic domains. Sequence alignments also have revealed linker regions between the catalytic domains and at the Nand C-termini of individual polypeptides. The sequences of these linker regions are less well conserved than are those for the catalytic domains, which is in part how linker regions are identified. Linker regions can be important for proper association between domains and between the individual polypeptides that comprise the PKS complex. One can thus view the linkers and domains together as creating a scaffold on which the domains and modules are positioned in the correct orientation to be active. This organization and positioning, if retained, permits PKS domains of different or identical substrate specificities to be substituted (usually at the DNA level) between PKS enzymes by various available methodologies. In selecting the boundaries of, for example, an AT replacement, one can thus make the replacement so as to retain the linkers of the recipient PKS or to replace them with the linkers of the donor PKS AT domain, or, preferably, make both constructs to ensure that the correct linker regions between the KS and AT domains have been included in at least one of the engineered enzymes. Thus, there is considerable flexibility in the design of new PKS enzymes with the result that known polyketides can be produced more effectively, and novel polyketides useful as pharmaceuticals or for other purposes can be made.

By appropriate application of recombinant DNA technology, a wide variety of polyketides can be prepared in a variety of different host cells provided one has access to nucleic acid compounds that encode PKS proteins and polyketide modification enzymes. The present invention helps meet the need for such nucleic acid compounds by providing recombinant vectors that encode the FK-520 PKS enzyme and various FK-520 modification enzymes. Moreover, while the FK-506 and FK-520 polyketides have many useful activities, there remains a need for compounds with similar useful activities but with better pharmacokinetic profile and metabolism and fewer side-effects. The present invention helps meet the need for such compounds as well.

Summary of the Invention

In one enrodiment, the present invention provides recombinant DNA vectors that encode all or part of the FK-520 PKS enzyme. Illustrative vectors of the invention include cosmid pKOS034-120, pKOS034-124, pKOS065-C31, pKOS065-C3,

pKOS065-M27, and pKOS065-M21. The invention also provides nucleic acid compounds that encode the various domains of the FK-520 PKS, i.e., the KS, AT, ACP, KR, DH, and ER domains. These compounds can be readily used, alone or in combination with nucleic acids encoding other FK-520 or non-FK-520 PKS domains, as intermediates in the construction of recombinant vectors that encode all or part of PKS enzymes that make novel polyketides.

5

10

15

20

25

30

The invention also provides isolated nucleic acids that encode all or part of one or more modules of the FK-520 PKS, each module comprising a ketosynthase activity, an acyl transferase activity, and an acyl carrier protein activity. The invention provides an isolated nucleic acid that encodes one or more open reading frames of FK-520 PKS genes, said open reading frames comprising coding sequences for a CoA ligase activity, an NRPS activity, or two or more extender modules. The invention also provides recombinant expression vectors containing these nucleic acids.

In another embodiment, the invention provides isolated nucleic acids that encode all or a part of a PKS that contains at least one module in which at least one of the domains in the module is a domain from a non-FK-520 PKS and at least one domain is from the FK-520 PKS. The non-FK-520 PKS domain or module originates from the rapamycin PKS, the FK-506 PKS, DEBS, or another PKS. The invention also provides recombinant expression vectors containing these nucleic acids.

In another embodiment, the invention provides a method of preparing a polyketide, said method comprising transforming a host cell with a recombinant DNA vector that encodes at least one module of a PKS, said module comprising at least one FK-520 PKS domain, and culturing said host cell under conditions such that said PKS is produced and catalyzes synthesis of said polyketide. In one aspect, the method is practiced with a *Streptomyces* host cell. In another aspect, the polyketide produced is FK-520. In another aspect, the polyketide produced is a polyketide related in structure to FK-520. In another aspect, the polyketide produced is a polyketide related in structure to FK-506 or rapamycin.

In another embodiment, the invention provides a set of genes in recombinant form sufficient for the synthesis of ethylmalonyl CoA in a heterologous host cell. These genes and the methods of the invention enable one to create recombinant host cells with the ability to produce polyketides or other compounds that require ethylmalonyl CoA for biosynthesis. The invention also provides recombinant nucleic acids that encode AT domains specific for ethylmalonyl CoA. Thus, the compounds of the invention can be

used to produce polyketides requiring ethylmalonyl CoA in host cells that otherwise are unable to produce such polyketides.

In another embodiment, the invention provides a set of genes in recombinant form sufficient for the synthesis of 2-hydroxymalonyl CoA and 2-methoxymalonyl CoA in a heterologous host cell. These genes and the methods of the invention enable one to create recombinant host cells with the ability to produce polyketides or other compounds that require 2-hydroxymalonyl CoA for biosynthesis. The invention also provides recombinant nucleic acids that encode AT domains specific for 2-hydroxymalonyl CoA and 2-methoxymalonyl CoA. Thus, the compounds of the invention can be used to produce polyketides requiring 2-hydroxymalonyl CoA or 2-methoxymalonyl CoA in host cells that are otherwise unable to produce such polyketides.

In another embodiment, the invention provides a compound related in structure to FK-520 or FK-506 that is useful in the treatment of a medical condition. These compounds include compounds in which the C-13 methoxy group is replaced by a moiety selected from the group consisting of hydrogen, methyl, and ethyl moieties. Such compounds are less susceptible to the main *in vivo* pathway of degradation for FK-520 and FK-506 and related compounds and thus exhibit an improved pharmacokinetic profile. The compounds of the invention also include compounds in which the C-15 methoxy group is replaced by a moiety selected from the group consisting of hydrogen, methyl, and ethyl moieties. The compounds of the invention also include the above compounds further modified by chemical methodology to produce derivatives such as, but not limited to, the C-18 hydroxyl derivatives, which have potent neurotrophin but not immunosuppression activities.

Thus, the invention provides polyketides having the structure:

5

10

15

20

25

wherein, R₁ is hydrogen, methyl, ethyl, or allyl; R₂ is hydrogen or hydroxyl, provided that when R₂ is hydrogen, there is a double bond between C-20 and C-19; R₃ is hydrogen

or hydroxyl; R₄ is methoxyl, hydrogen, methyl, or ethyl; and R₅ is methoxyl, hydrogen, methyl, or ethyl; but not including FK-506, FK-520, 18-hydroxy-FK-520, and 18-hydroxy-FK-506. The invention provides these compounds in purified form and in pharmaceutical compositions.

In another embodiment, the invention provides a method for treating a medical condition by administering a pharmaceutically efficacious dose of a compound of the invention. The compounds of the invention may be administered to achieve immunosuppression or to stimulate nerve growth and regeneration.

5

10

15

20

25

30

35

These and other embodiments and aspects of the invention will be more fully understood after consideration of the attached Drawings and the abrief description below, together with the detailed description, examples, and claims that follow.

Brief Description of the Drawings

Figure 1 shows a diagram of the FK-520 biosynthetic gene cluster. The top line provides a scale in kilobase pairs (kb). The second line shows a restriction map with selected restriction enzyme recognition sequences indicated. K is *KpnI*; X is *XhoI*, S is *SacI*; P is *PstI*; and E is *EcoRI*. The third line indicates the position of FK-520 PKS and related genes. Genes are abbreviated with a one letter designation, i.e., C is *fkbC*. Immediately under the third line are numbered segments showing where the loading module (L) and ten different extender modules (numbered 1 - 10) are encoded on the various genes shown. At the bottom of the Figure, the DNA inserts of various cosmids of the invention (i.e., 34-124 is cosmid pKOS034-124) are shown in alignment with the FK-520 biosynthetic gene cluster.

Figure 2 shows the loading module (load), the ten extender modules, and the peptide synthetase domain of the FK-520 PKS, together with, on the top line, the genes that encode the various domains and modules. Also shown are the various intermediates in FK-520 biosynthesis, as well as the structure of FK-520, with carbons 13, 15, 21, and 31 numbered. The various domains of each module and subdomains of the loading module are also shown. The darkened circles showing the DH domains in modules 2, 3, and 4 indicate that the dehydratase domain is not functional as a dehydratase; this domain may affect the stereochemistry at the corresponding position in the polyketide. The substituents on the FK-520 structure that result from the action of non-PKS enzymes are also indicated by arrows, together with the types of enzymes or the genes that code for the enzymes that mediate the action. Although the methyltransferase is shown acting at the C-13 and C-15 hydroxyl groups after release of the polyketide from the PKS, the

methyltransferase may act on the 2-hydroxymalonyl substrate prior to or contemporaneously with its incorporation during polyketide synthesis.

5

10

15

20

25

30

35

Figure 3 shows a close-up view of the left end of the FK-520 gene cluster, which contains at least ten additional genes. The ethyl side chain on carbon 21 of FK-520 (Figure 2) is derived from an ethylmalonyl CoA extender unit that is incorporated by an ethylmalonyl specific AT domain in extender module 4 of the PKS. At least four of the genes in this region code for enzymes involved in ethylmalonyl biosynthesis. The polyhydroxybutyrate depolymerase is involved in maintaining hydroxybutyryl-CoA pools during FK-520 production. Polyhydroxybutyrate accumulates during vegetative growth and disappears during stationary phase in other *Streptomyces* (Ranade and Vining, 1993, *Can. J. Microbiol. 39*:377). Open reading frames with unknown function are indicated with a question mark.

Figure 4 shows a biosynthetic pathway for the biosynthesis of ethylmalonyl CoA from acetoacetyl CoA consistent with the function assigned to four of the genes in the FK-520 gene cluster shown in Figure 3.

Figure 5 shows a close-up view of the right-end of the FK-520 PKS gene cluster (and of the sequences on cosmid pKOS065-C31). The genes shown include *fkbD*, *fkbM* (a methyl transferase that methylates the hydroxyl group on C-31 of FK-520), *fkbN* (a homolog of a gene described as a regulator of cholesterol oxidase and that is believed to be a transcriptional activator), *fkbQ* (a type II thioesterase, which can increase polyketide production levels), and *fkbS* (a crotonyl-CoA reductase involved in the biosynthesis of ethylmalonyl CoA).

Figure 6 shows the proposed degradative pathway for tacrolimus (FK-506) metabolism.

Figure 7 shows a schematic process for the construction of recombinant PKS genes of the invention that encode PKS enzymes that produce 13-desmethoxy FK-506 and FK-520 polyketides of the invention, as described in Example 4, below.

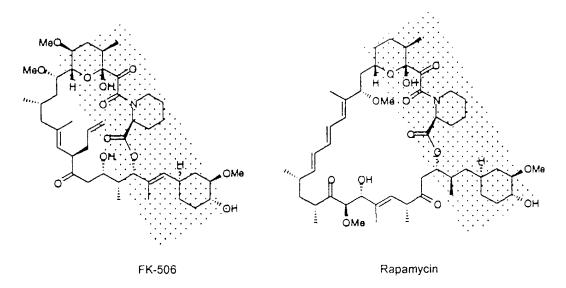
Figure 8, in Parts A and B, shows certain compounds of the invention preferred for dermal application in Part A and a synthetic route for making those compounds in Part B.

Detailed Description of the Invention

Given the valuable pharmaceutical properties of polyketides, there is a need for methods and reagents for producing large quantities of polyketides, as well as for producing related compounds not found in nature. The present invention provides such

methods and reagents, with particular application to methods and reagents for producing the polyketides known as FK-520, also known as ascomycin or L-683,590 (see Holt *et al.*, 1993, *JACS 115*:9925), and FK-506, also known as tacrolimus. Tacrolimus is a macrolide immunosuppressant used to prevent or treat rejection of transplanted heart, kidney, liver, lung, pancreas, and small bowel allografts. The drug is also useful for the prevention and treatment of graft-versus-host disease in patients receiving bone marrow transplants, and for the treatment of severe, refractory uveitis. There have been additional reports of the unapproved use of tacrolimus for other conditions, including alopecia universalis, autoimmune chronic active hepatitis, inflammatory bowel disease, multiple sclerosis, primary biliary cirrhosis, and scleroderma. The invention provides methods and reagents for making novel polyketides related in structure to FK-520 and FK-506, and structurally related polyketides such as rapamycin.

The FK-506 and rapamycin polyketides are potent immunosuppressants, with chemical structures shown below.



15

20

25

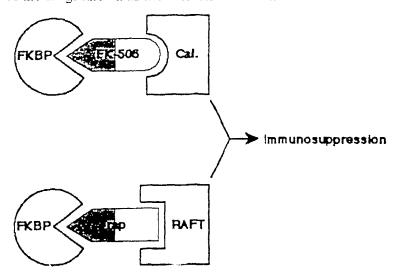
5

10

FK-520 differs from FK-506 in that it lacks the allyl group at C-21 of FK-506, having instead an ethyl group at that position, and has similar activity to FK-506, albeit reduced immunosuppressive activity.

These compounds act through initial formation of an intermediate complex with protein "immunophilins" known as FKBPs (FK-506 binding proteins), including FKBP-12. Immunophilins are a class of cytosolic proteins that form complexes with molecules such as FK-506, FK-520, and rapamycin that in turn serve as ligands for other cellular targets involved in signal transduction. Binding of FK-506, FK-520, and rapamycin to FKBP occurs through the structurally similar segments of the polyketide molecules, known as the "FKBP-binding domain" (as generally but not precisely indicated by the

stippled regions in the structures above). The FK-506-FKBP complex then binds calcineurin, while the rapamycin-FKBP complex binds to a protein known as RAFT-1. Binding of the FKBP-polyketide complex to these second proteins occurs through the dissimilar regions of the drugs known as the "effector" domains.



5

10

The three component FKBP-polyketide-effector complex is required for signal transduction and subsequent immunosuppressive activity of FK-506, FK-520, and rapamycin. Modifications in the effector domains of FK-506, FK-520, and rapamycin that destroy binding to the effector proteins (calcineurin or RAFT) lead to loss of immunosuppressive activity, even though FKBP binding is unaffected. Further, such analogs antagonize the immunosuppressive effects of the parent polyketides, because they compete for FKBP. Such non-immunosuppressive analogs also show reduced toxicity (see Dumont et al., 1992, Journal of Experimental Medicine 176, 751-760), indicating that much of the toxicity of these drugs is not linked to FKBP binding.

15

20

25

In addition to immunosuppressive activity, FK-520, FK-506, and rapamycin have neurotrophic activity. In the central nervous system and in peripheral nerves, immunophilins are referred to as "neuroimmunophilins". The neuroimmunophilin FKBP is markedly enriched in the central nervous system and in peripheral nerves. Molecules that bind to the neuroimmunophilin FKBP, such as FK-506 and FK-520, have the remarkable effect of stimulating nerve growth. *In vitro*, they act as neurotrophins, i.e., they promote neurite outgrowth in NGF-treated PC12 cells and in sensory neuronal cultures, and in intact animals, they promote regrowth of damaged facial and sciatic nerves, and repair lesioned serotonin and dopamine neurons in the brain. See Gold *et al.*, Jun. 1999, *J. Pharm. Exp. Ther.* 289(3): 1202-1210; Lyons *et al.*, 1994, *Proc. National Academy of Science* 91: 3191-3195; Gold *et al.*, 1995, *Journal of Neuroscience* 15:

7509-7516; and Steiner et al., 1997, Proc. National Academy of Science 94: 2019-2024. Further, the restored central and peripheral neurons appear to be functional.

5

10

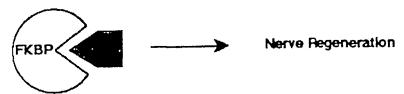
15

20

25

Compared to protein neurotrophic molecules (BNDF, NGF, etc.), the small-molecule neurotrophins such as FK-506, FK-520, and rapamycin have different, and often advantageous, properties. First, whereas protein neurotrophins are difficult to deliver to their intended site of action and may require intra-cranial injection, the small-molecule neurotrophins display excellent bioavailability; they are active when administered subcutaneously and orally. Second, whereas protein neurotrophins show quite specific effects, the small-molecule neurotrophins show rather broad effects. Finally, whereas protein neurotrophins often show affects on normal sensory nerves, the small-molecule neurotrophins do not induce aberrant sprouting of normal neuronal processes and seem to affect damaged nerves specifically. Neuroimmunophilin ligands have potential therapeutic utility in a variety of disorders involving nerve degeneration (e.g. multiple sclerosis, Parkinson's disease, Alzheimer's disease, stroke, traumatic spinal cord and brain injury, peripheral neuropathies).

Recent studies have shown that the immunosuppressive and neurite outgrowth activity of FK-506, FK-520, and rapamycin can be separated; the neuroregenerative activity in the absence of immunosuppressive activity is retained by agents which bind to FKBP but not to the effector proteins calcineurin or RAFT. See Steiner *et al.*, 1997, *Nature Medicine 3*: 421-428.



Available structure-activity data show that the important features for neurotrophic activity of rapamycin, FK-520, and FK-506 lie within the common, contiguous segments of the macrolide ring that bind to FKBP. This portion of the molecule is termed the "FKBP binding domain" (see VanDuyne *et al.*, 1993, *Journal of Molecular Biology 229*: 105-124.). Nevertheless, the effector domains of the parent macrolides contribute to conformational rigidity of the binding domain and thus indirectly contribute to FKBP binding.

"FKBP binding domain"

There are a number of other reported analogs of FK-506, FK-520, and rapamycin that bind to FKBP but not the effector protein calcineum or RAFT. These analogs show effects on nerve regeneration without immunosuppressive effects.

Naturally occurring FK-520 and FK-506 analogs include the antascomycins, which are FK-506-like macrolides that lack the functional groups of FK-506 that bind to calcineurin (see Fehr *et al.*, 1996, *The Journal of Antibiotics* 49: 230-233). These molecules bind FKBP as effectively as does FK-506; they antagonize the effects of both FK-506 and rapamycin, yet lack immunosuppressive activity.

Antascomycin A

10

15

5

Other analogs can be produced by chemically modifying FK-506, FK-520, or rapamycin. One approach to obtaining neuroimmunophilin ligands is to destroy the effector binding region of FK-506, FK-520, or rapamycin by chemical modification. While the chemical modifications permitted on the parent compounds are quite limited, some useful chemically modified analogs exist. The FK-520 analog L-685,818 (ED₅₀ = 0.7 nM for FKBP binding; see Dumont *et al.*, 1992), and the rapamycin analog WAY-124,466 (IC₅₀ = 12.5 nM; see Ocain *et al.*, 1993, *Biochemistry Biophysical Research Communications 192*: 1340-134693) are about as effective as FK-506, FK-520, and rapamycin at promoting neurite outgrowth in sensory neurons (see Steiner *et al.*, 1997).

One of the few positions of rapamycin that is readily amenable to chemical modification is the allylic 16-methoxy group; this reactive group is readily exchanged by acid-catalyzed nucleophilic substitution. Replacement of the 16-methoxy group of rapamycin with a variety of bulky groups has produced analogs showing selective loss of immunosuppressive activity while retaining FKBP-binding (see Luengo *et al.*, 1995, *Chemistry & Biology 2*: 471-481). One of the best compounds, 1, below, shows complete loss of activity in the splenocyte proliferation assay with only a 10-fold reduction in binding to FKBP.

5

10

15

There are also synthetic analogs of FKBP binding domains. These compounds reflect an approach to obtaining neuroimmunophilin ligands based on "rationally designed" molecules that retain the FKBP-binding region in an appropriate conformation for binding to FKBP, but do not possess the effector binding regions. In one example, the ends of the FKBP binding domain were tethered by hydrocarbon chains (see Holt *et al.*, 1993, *Journal of the American Chemical Society 115*: 9925-9938); the best analog, 2, below, binds to FKBP about as well as FK-506. In a similar approach, the ends of the

FKBP binding domain were tethered by a tripeptide to give analog 3, below, which binds

to FKBP about 20-fold poorer than FK-506. These compounds are anticipated to have neuroimmunophilin binding activity.

In a primate MPTP model of Parkinson's disease, administration of FKBP ligand GPI-1046 caused brain cells to regenerate and behavioral measures to improve. MPTP is a neurotoxin, which, when administered to animals, selectively damages nigral-striatal dopamine neurons in the brain, mimicking the damage caused by Parkinson's disease. Whereas, before treatment, animals were unable to use affected limbs, the FKBP ligand restored the ability of animals to feed themselves and gave improvements in measures of locomotor activity, neurological outcome, and fine motor control. There were also corresponding increases in regrowth of damaged nerve terminals. These results demonstrate the utility of FKBP ligands for treatment of diseases of the CNS.

5

10

15

20

25

30

From the above description, two general approaches towards the design of non-immunosuppressant, neuroimmunophilin ligands can be seen. The first involves the construction of constrained cyclic analogs of FK-506 in which the FKBP binding domain is fixed in a conformation optimal for binding to FKBP. The advantages of this approach are that the conformation of the analogs can be accurately modeled and predicted by computational methods, and the analogs closely resemble parent molecules that have proven pharmacological properties. A disadvantage is that the difficult chemistry limits the numbers and types of compounds that can be prepared. The second approach involves the trial and error construction of acyclic analogs of the FKBP binding domain by conventional medicinal chemistry. The advantages to this approach are that the chemistry is suitable for production of the numerous compounds needed for such interactive chemistry-bioassay approaches. The disadvantages are that the molecular types of compounds that have emerged have no known history of appropriate pharmacological properties, have rather labile ester functional groups, and are too conformationally mobile to allow accurate prediction of conformational properties.

The present invention provides useful methods and reagents related to the first approach, but with significant advantages. The invention provides recombinant PKS

genes that produce a wide variety of polyketides that cannot otherwise be readily synthesized by chemical methodology alone. Moreover, the present invention provides polyketides that have either or both of the desired immunosuppressive and neurotrophic activities, some of which are produced only by fermentation and others of which are produced by fermentation and chemical modification. Thus, in one aspect, the invention provides compounds that optimally bind to FKBP but do not bind to the effector proteins. The methods and reagents of the invention can be used to prepare numerous constrained cyclic analogs of FK-520 in which the FKBP binding domain is fixed in a conformation optimal for binding to FKBP. Such compounds will show neuroimmunophilin binding (neurotrophic) but not immunosuppressive effects. The invention also allows direct manipulation of FK-520 and related chemical structures via genetic engineering of the enzymes involved in the biosynthesis of FK-520 (as well as related compounds, such as FK-506 and rapamycin); similar chemical modifications are simply not possible because of the complexity of the structures. The invention can also be used to introduce "chemical handles" into normally inert positions that permit subsequent chemical modifications.

5

10

15

20

25

30

Several general approaches to achieve the development of novel neuroimmunophilin ligands are facilitated by the methods and reagents of the present invention. One approach is to make "point mutations" of the functional groups of the parent FK-520 structure that bind to the effector molecules to eliminate their binding potential. These types of structural modifications are difficult to perform by chemical modification, but can be readily accomplished with the methods and reagents of the invention.

A second, more extensive approach facilitated by the present invention is to utilize molecular modeling to predict optimal structures *ab initio* that bind to FKBP but not effector molecules. Using the available X-ray crystal structure of FK-520 (or FK-506) bound to FKBP, molecular modeling can be used to predict polyketides that should optimally bind to FKBP but not calcineurin. Various macrolide structures can be generated by linking the ends of the FKBP-binding domain with "all possible" polyketide chains of variable length and substitution patterns that can be prepared by genetic manipulation of the FK-520 or FK-506 PKS gene cluster in accordance with the methods of the invention. The ground state conformations of the virtual library can be determined, and compounds that possess binding domains most likely to bind well to FKBP can be prepared and tested.

Once a compound is identified in accordance with the above approaches, the invention can be used to generate a focused library of analogs around the lead candidate, to "fine tune" the compound for optimal properties. Finally, the genetic engineering methods of the invention can be directed towards producing "chemical handles" that enable medicinal chemists to modify positions of the molecule previously inert to chemical modification. This opens the path to previously prohibited chemical optimization of lead compounds by time-proven approaches.

Moreover, the present invention provides polyketide compounds and the recombinant genes for the PKS enzymes that produce the compounds that have significant advantages over FK-506 and FK-520 and their analogs. The metabolium and pharmacokinetics of tacrolimus has been exaministely studied, and FK-520 is believed to be similar in these respects. Absorption of tacrolimus is rapid, variable, and incomplete from the gastrointestinal tract (Harrison's Principles of Internal Medicine, 14th edition, 1998, McGraw Hill, 14, 20, 21, 64-67). The mean bioavailability of the oral dosage form is 27%. (range 5 to 65%). The volume of distribution (VolD) based on plasma is 5 to 65 L per kg of body weight (L/kg), and is much higher than the VolD based on whole blood concentrations, the difference reflecting the binding of tacrolimus to red blood cells. Whole blood concentrations may be 12 to 67 times the plasma concentrations. Protein binding is high (75 to 99%), primarily to albumin and alpha1-acid glycoprotein. The half-life for distribution is 0.9 hour; elimination is biphasic and variable: terminal-11.3 hr (range, 3.5 to 40.5 hours). The time to peak concentration is 0.5 to 4 hours after oral administration.

Tacrolimus is metabolized primarily by cytochrome P450 3A enzymes in the liver and small intestine. The drug is extensively metabolized with less than 1% excreted unchanged in urine. Because hepatic dysfunction decreases clearance of tacrolimus, doses have to be reduced substantially in primary graft non-function, especially in children. In addition, drugs that induce the cytochrome P450 3A enzymes reduce tacrolimus levels, while drugs that inhibit these P450s increase tacrolimus levels.

Tacrolimus bioavailability doubles with co-administration of ketoconazole, a drug that inhibits P450 3A. See, Vincent *et al.*, 1992, *In vitro* metabolism of FK-506 in rat, rabbit, and human liver microsomes: Identification of a major metabolite and of cytochrome P450 3A as the major enzymes responsible for its metabolism, *Arch. Biochem. Biophys.* 294: 454-460; Iwasaki *et al.*, 1993, Isolation, identification, and biological activities of oxidative metabolites of FK-506, a potent immunosuppressive macrolide lactone, *Drug Metabolism & Disposition 21*: 971-977; Shiraga *et al.*, 1994, Metabolism of FK-506, a

potent immunosuppressive agent, by cytochrome P450 3A enzymes in rat, dog, and human liver microsomes, *Biochem. Pharmacol.* 47: 727-735; and Iwasaki et al., 1995, Further metabolism of FK-506 (Tacrolimus); Identification and biological activities of the metabolites oxidized at multiple sites of FK-506. *Drug Metabolism & Disposition* 23: 28-34. The cytochrome P450 3A subfamily of isozymes has been implicated as important in this degradative process.

Structures of the eight isolated metabolites formed by liver microsomes are shown in Figure 6. Four metabolites of FK-506 involve demethylation of the oxygens on carbons 13, 15, and 31, and hydroxylation of carbon 12. The 13-demethylated (hydroxy) compounds undergo cycnizations of the 13-hydroxy at C-10 to give MI, MVI and MVII, and the 12-hydroxy metabolite at C-10 to give I. Another four metabolites formed by oxidation of the four metabolites mentioned above were isolated by liver microsomes from dexamethasone treated rats. Three of these are metabolites doubly demethylated at the methoxy groups on carbons 15 and 31 (M-V), 13 and 31 (M-VI), and 13 and 15 (M-VII). The fourth, M-VIII, was the metabolite produced after demethylation of the 31-methoxy group, followed by formation of a fused ring system by further oxidation. Among the eight metabolites, M-II has immunosuppressive activity comparable to that of FK-506, whereas the other metabolites exhibit weak or negligible activities. Importantly, the major metabolite of human, dog, and rat liver microsomes is the 13-demethylated and cyclized FK-506 (M-I).

Thus, the major metabolism of FK-506 proceeds via 13-demethylation followed by cyclization to the inactive M-I, this representing about 90% of the metabolic products after a 10 minute incubation with liver microsomes. Analogs of tacrolimus that do not possess a C-13 methoxy group would not be susceptible to the first and most important biotransformation in the destructive metabolism of tacrolimus (i.e. cyclization of 13-hydroxy to C-10). Thus, a 13-desmethoxy analog of FK-506 should have a longer half-life in the body than does FK-506. The C-13 methoxy group is believed not to be required for binding to FKBP or calcineurin. The C-13 methoxy is not present on the identical position of rapamycin, which binds to FKBP with equipotent affinity as tacrolimus. Also, analysis of the 3-dimensional structure of the FKBP-tacrolimus-calcineurin complex shows that the C-13 methoxy has no interaction with FKBP and only a minor interaction with calcineurin. The present invention provides C-13-desmethoxy analogs of FK-506 and FK-520, as well as the recombinant genes that encode the PKS enzymes that catalyze their synthesis and host cells that produce the compounds.

These compounds exhibit, relative to their naturally occurring counterparts, prolonged immunosuppressive action *in vivo*, thereby allowing a lower dosage and/or reduced frequency of administration. Dosing is more predictable, because the variability in FK-506 dosage is largely due to variation of metabolism rate. FK-506 levels in blood can vary widely depending on interactions with drugs that induce or innibit cytochrome P450 3A (summarized in USP Drug Information for the Health Care Professional). Of particular importance are the numerous drugs that inhibit or compete for CYP 3A, because they increase FK-506 blood levels and lead to toxicity (Prograf package insert, Fujisawallus, Rev 4/97, Rec 6/97). Also important are the drugs that induce P450 3A (e.g. Dexardone), because they decrease FK-506 blood levels and reduce efficacy. Because the major site of CYP 3A action on FK-506 is removed in the analogs provided by the present invention, those analogs are not as susceptible to drug interactions as the naturally occurring compounds.

Hyperglycemia, nephrotoxicity, and neurotoxicity are the most significant
adverse effects resulting from the use of FK-506 and are believed to be similar for FK520. Because these effects appear to occur primarily by the same mechanism as the
immunosuppressive action (i.e. FKBP-calcineurin interaction), the intrinsic toxicity of
the desmethoxy analogs may be similar to FK-506. However, toxicity of FK-506 is dose
related and correlates with high blood levels of the drug (Prograf package insert,
FujisawaCUS, Rev 4/97, Rec 6/97). Because the levels of the compounds provided by
the present invention should be more controllable, the incidence of toxicity should be
significantly decreased with the 13-desmethoxy analogs. Some reports show that certain
FK-506 metabolites are more toxic than FK-506 itself, and this provides an additional
reason to expect that a CYP 3A resistant analog can have lower toxicity and a higher
therapeutic index.

Thus, the present invention provides novel compounds related in structure to FK-506 and FK-520 but with improved properties. The invention also provides methods for making these compounds by fermentation of recombinant host cells, as well as the recombinant host cells, the recombinant vectors in those host cells, and the recombinant proteins encoded by those vectors. The present invention also provides other valuable materials useful in the construction of these recombinant vectors that have many other important applications as well. In particular, the present invention provides the FK-520 PKS genes, as well as certain genes involved in the biosynthesis of FK-520 in recombinant form.

30

FK-520 is produced at relatively low levels in the naturally occurring cells, *Streptomyces hygroscopicus* var. *ascomyceticus*, in which it was first identified. Thus, another benefit provided by the recombinant FK-520 PKS and related genes of the present invention is the ability to produce FK-520 in greater quantities in the recombinant host cells provided by the invention. The invention also provides methods for making novel FK-520 analogs, in addition to the desmethoxy analogs described above, and derivatives in recombinant host cells of any origin.

The biosynthesis of FK-520 involves the action of several enzymes. The FK-520 PKS enzyme, which is composed of the *fkbA*, *fkbB*, *fkbC*, and *fkbP* gene products, synthesizes the core structure of the molecule. There is also a hydroxylation at C-9 mediated by the P450 hydroxylase that is the *fkbD* gene product and that is oxidized by the *fkbO* gene product to result in the formation of a keto group at C-9. There is also a methylation at C-31 that is mediated by an O-methyltransferase that is the *fkbM* gene product. There are also methylations at the C-13 and C-15 positions by a methyltransferase believed to be encoded by the fkbG gene; this methyltransferase may act on the hydroxymalonyl CoA substrates prior to binding of the substrate to the AT domains of the PKS during polyketide synthesis. The present invention provides the genes encoding these enzymes in recombinant form. The invention also provides the genes encoding the enzymes involved in ethylmalonyl CoA and 2-hydroxymalonyl CoA biosynthesis in recombinant form. Moreover, the invention provides *Streptomyces hygroscopicus* var. *ascomyceticus* recombinant host cells lacking one or more of these genes that are useful in the production of useful compounds.

The cells are useful in production in a variety of ways. First, certain cells make a useful FK-520-related compound merely as a result of inactivation of one or more of the FK-520 biosynthesis genes. Thus, by inactivating the C-31 O-methyltransferase gene in *Streptomyces hygroscopicus* var. *ascomyceticus*, one creates a host cell that makes a desmethyl (at C-31) derivative of FK-520. Second, other cells of the invention are unable to make FK-520 or FK-520 related compounds due to an inactivation of one or more of the PKS genes. These cells are useful in the production of other polyketides produced by PKS enzymes that are encoded on recombinant expression vectors and introduced into the host cell.

Moreover, if only one PKS gene is inactivated, the ability to produce FK-520 or an FK-520 derivative compound is restored by introduction of a recombinant expression vector that contains the functional gene in a modified or unmodified form. The introduced gene produces a gene product that, together with the other endogenous and

functional gene products, produces the desired compound. This methodology enables one to produce FK-520 derivative compounds without requiring that all of the genes for the PKS enzyme be present on one or more expression vectors. Additional applications and benefits of such cells and methodology will be readily apparent to those of skill in the art after consideration of how the recombinant genes were isolated and employed in the construction of the compounds of the invention.

The FK-520 biosynthetic genes were isolated by the following procedure. Genomic DNA was isolated from *Streptomyces hygroscopicus* var. *ascomyceticus* (ATCC 14891) using the lysozyme/proteinase K protocol described in Genetic Manipulation of *Streptomyces* - A Laboratory Manual (Hopwood *et al.*, 1986). The average size of the DNA was estimated to be between 80 - 120 kb by electrophoresis on 0.3% agarose gels. A library was constructed in the SuperCosTM vector according to the manufacturer's instructions and with the reagents provided in the commercially available kit (Stratagene). Briefly, 100 µg of genomic DNA was partially digested with 4 units of *Sau*3A I for 20 min. in a reaction volume of 1 mL, and the fragments were dephosphorylated and ligated to SuperCos vector arms. The ligated DNA was packaged and used to infect log-stage XL1-BlueMR cells. A library of about 10,000 independent cosmid clones was obtained.

Based on recently published sequence from the FK-506 cluster (Motamedi and Shafiee, 1998, *Eur. J. Biochem. 256*: 528), a probe for the *fkbO* gene was isolated from ATCC 14891 using PCR with degenerate primers. With this probe, a cosmid designated pKOS034-124 was isolated from the library. With probes made from the ends of cosmid pKOS034-124, an additional cosmid designated pKOS034-120 was isolated. These cosmids (pKOS034-124 and pKOS034-120) were shown to contain DNA inserts that overlap with one another. Initial sequence data from these two cosmids generated sequences similar to sequences from the FK-506 and rapamycin clusters, indicating that the inserts were from the FK-520 PKS gene cluster. Two *Eco*RI fragments were subcloned from cosmids pKOS034-124 and pKOS034-120. These subclones were used to prepare shotgun libraries by partial digestion with *Sau*3AI, gel purification of fragments between 1.5 kb and 3 kb in size, and ligation into the pLitmus28 vector (New England Biolabs). These libraries were sequenced using dye terminators on a Beckmann CEQ2000 capillary electrophoresis sequencer, according to the manufacturer's protocols.

To obtain cosmids containing sequence on the left and right sides of the sequenced region described above, a new cosmid library of ATCC 14891 DNA was prepared essentially as described above. This new library was screened with a new fkbM

probe isolated using DNA from ATCC 14891. A probe representing the *fkbP* gene at the end of cosmid pKOS034-124 was also used. Several additional cosmids to the right of the previously sequenced region were identified. Cosmids pKOS065-C31 and pKOS065-C3 were identified and then mapped with restriction enzymes. Initial sequences from these cosmids were consistent with the expected organization of the cluster in this region. More extensive sequencing showed that both cosmids contained in addition to the desired sequences, other sequences not contiguous to the desired sequences on the host cell chromosomal DNA. Probing of additional cosmid libraries identified two additional cosmids, pKOS065-M27 and pKOS065-M21, that contained the desired sequences in a contiguous segment of chromosomal DNA. Cosmids pKOS034-124, pKOS034-120, pKOS065-M27, and pKOS065-M21 have been deposited with the American Type Culture Collection, Manassas, VA, USA. The complete nucleotide sequence of the coding sequences of the genes that encode the proteins of the FK-520 PKS are shown below but can also be determined from the cosmids of the invention deposited with the ATCC using standard methodology.

Referring to Figures 1 and 3, the FK-520 PKS gene cluster is composed of four open reading frames designated fkbB, fkbC, fkbA, and fkbP. The fkbB open reading frame encodes the loading module and the first four extender modules of the PKS. The fkbC open reading frame encodes extender modules five and six of the PKS. The fkbA open reading frame encodes extender modules seven, eight, nine, and ten of the PKS. The fkbP open reading frame encodes the NRPS of the PKS. Each of these genes can be isolated from the cosmids of the invention described above. The DNA sequences of these genes are provided below preceded by the following table identifying the start and stop codons of the open reading frames of each gene and the modules and domains contained therein.

Nucleotides (412 1026)	Gene or Domain
•	fkbW
complement (2020 - 3579)	fkbV
complement (3969 - 4496)	fkbR2
complement (4595 - 5488)	fkbR I
5601 - 6818	fkbE
6808 - 8052	fkbF
8156 - 8824	fkbG
complement (9122 - 9883)	fkbH
complement (9894 - 10994)	fkbI
complement (10987 - 11247)	fkbJ
complement (11244 - 12092)	fkbK
complement (12113 - 13150)	fkbL
complement (13212 - 23988)	fkbC
	complement (412 - 1836) complement (2020 - 3579) complement (3969 - 4496) complement (4595 - 5488) 5601 - 6818 6808 - 8052 8156 - 8824 complement (9122 - 9883) complement (9894 - 10994) complement (10987 - 11247) complement (11244 - 12092) complement (12113 - 13150)

	complement (23992 - 46573)	fkbB
	46754 - 47788	fkbO
	47785 - 52272	fkbP
	52275 - 71465	fkbA
5	71462 - 72628	fkbD
	72625 - 73407	fkbM
	complement (73460 - 76202)	fkbN
	complement (76336 - 77080)	fkbQ
	complement (77076 - 77535)	fkbS
10	complement (44974 - 46573)	CoA ligase of loading domain
	complement (43777 - 44629)	ER of loading domain
	complement (43144 - 43660)	ACP of loading domain
	complement (41842 - 43093)	KS of extender module 1 (KS1)
	complement(40609 - 41842)	AT1
15	complement (39442 - 40609)	DH1
	complement (38677 - 39307)	KR1
	complement (38371 - 38581)	ACP1
	complement (37145 - 38296)	KS2
	complement (35749 - 37144)	AT2
20	complement (34606 - 35749)	DH2 (inactive)
	complement (33823 - 34480)	KR2
	complement (33505 - 33715)	ACP2
	complement (32185 - 33439)	KS3
	complement (31018 - 32185)	AT3
25	complement (29869 - 31018)	DH3 (inactive)
	complement (29092 - 29740)	KR3
	complement (28750 - 28960)	ACP3
	complement (27430 - 28684)	KS4
	complement (26146 - 27430)	AT4
30	complement (24997 - 26146)	DH4 (inactive)
	complement (24163 - 24373)	ACP4
	complement (22653 - 23892)	KS5
	complement (21420 - 22653)	AT5
	complement (20241 - 21420)	DH5
35	complement (19464 - 20097)	KR5
	complement (19116 - 19326)	ACP5
	complement (17820 - 19053)	KS6
	complement (16587 - 17820)	AT6
	complement (15438 - 16587)	DH6
40	complement (14517 - 15294)	ER6
	complement (13761 - 14394)	KR6
	complement (13452 - 13662)	ACP6
	5236? - 53576	KS7
4.5	53577 - 54716	AT7
45	54717 - 55871	DH7
	56019 - 56819	ER7
	56943 - 57575	KR7
	57710 - 57920 57000 - 50242	ACP7
50	57990 - 59243 50244 - 60208	KS8
50	59244 - 60398	AT8
	60399 - 61412	DH8 (inactive)
	61548 - 62180	KR8

		- 62537 - 63854		ACP8 KS9			
		- 65084		AT9			
		- 66254		DH9			
5		- 67175		ER9			
	67299	- 67931		KR9			
	68094	- 68303		ACP9			
	68397	- 69653		KS10			
	69654	- 70985		AT10			
10	71064	- 71273		ACP10			
	-		ATGAAGTCCT		000003A0010	GTGAACACCT	CGCCGCTGCT
	1 h 1 d 1	TOTACGGACC TTACAAGATC	ACTICAGICA CICACATIGO	GOGGCGATTG		TUMTUUGGAA	TAAAGGGCGG GAGGCAAACC
15	181	GAAAGGGCGC	GGGGGGTCCG	CACCACCACCA	.400000	DSAGAGTGGC	GCACCCGCGC
• •	241	ACCGTCACCT	CTCTCCCCCG	CCGGCGGGAT	3000330303	ACGGTTGG	GCTCTC(CCG
	301	ACGCTGAACA	CCCGCGCGGT	GTGGCGTCGG	GGACACCSCC	TGGCATCGGC	CGGGTGACGG
	361	TACGGGGAGG	GCGTACGGCG	GCCGTGGCTC	GTGCTCACCC	coscoggggg	GTCATCCGTC
20	421	GAGACGGCAC	TOGGOGAGCA	GGGACGCCTG	GTCGGCACCT		CGACCGTGTG
20	481	GTTCGCGGGC	GGGCGGTGGC AGCAMAGGCC	GGGTGGTGAG GGAGTCGGTC	3333123373	AGGGCGGTGA TOGACGAGGG	AGGCTGAGCG CGTCGGTGTG
	601	GREACACGGC GGTGCCGTCC	TOGATOOGGT	AGTAGOGGTA		1300360TGCC	GGACATACGC
	661	GOGTACACGT	CGGAGCCGGG	GOGGOAGGOA	GCAGCACCTC	GAGAGTGCCT	GGATGGTGAT
	721	CAGCGGCTTG	CCGATACGAC	CGGTCAACGC	GATOCGTTTCC	ADGGCGCGT	GGACGCUGGA
25	781	GGAGCGGGTG	GCGTAGTCGT	AGTCGGCATC	GCAGCCGGG	Accendeded	GGGCGCAATA
	841	CGGTGTGCCG	GCTTCCTTCT	CCCCATCGAA	GCCGGGGTCG	AACTCCTCGC	GGTAGACGCG
	901 961	CTGCGTCAGA GAACCCGGCG	TCCCAGTAGA CGGAGCAGCG	CCTCGTGGTG	GTACGGCCAC CTGGCCGGCT	AAGAACTOGG GOGGGGGCOGC	AGTCGGCCGG CTGCCGCGTA
	1021	GGTGGGGTAG	TCGCGCAGGG	CGGCCGGCAG		AGGTTGGGAC	
30	1381	CCACAGGGTG	CCTTCCCAGT	CGACTCCTCC	GTCGTACAGC	TOGGGATGGT	TCTCCAGCTG
	114	CCAGCGCACG	AGGTAGCCGC	CGTTGGACAT	CCCGGGTGACC	AGGGTGCGCT	CGAGCGGCCG
	1201	GTGGTAGCGC		ACGCGCGGGC	GGCCCGGGTC	AGCTGGGTGA	
	1261	CCACTCGGCG	ACGGCGTCGC	CCGGCCGGGA	GCCATCACGG	TAGAACGCGG ACCCAGTCGG	GGCCGGTGTT
35	1321 1381	GCCCTTGTCG	GTGGCGGCGT TACTGCTCGC	AGGCGTAACC GGTTACCGGG	GGTGCCGGCC	ACGACCAGGC	CACCGTTCCA
27.0	1441	GCGGTCGGGC		CGAACTGGGC	GTCGTGGTTC		TGGTGTTGGT
	1501	GGTGGAGGTG	TCGGGGAAGT	AGCCGTCGAT	CTGGATCCCG	GGCACTCCGG	TGGGAGTGGC
	1561	CAGGTTCTTG	GGCGTCAGCC	CTGCCCAGTC		GTGTGGCCGG	
10		TCCCGCCGTG	GTCAGCTCGT	CCAGGCAGTC	GGCCTGCTGA	CGTGCCGCCG	CCGGGACACG
40	1681	CGGTGAGGGG	AGACGGGGCGC	AG FGACCGTC		GGAGCAGGCC	CCCTCCTGGC
	1801					AGCCTCCAGA	
	1861					ACATGGGTGC	
		ACTGAGGCCC					
45		GGCGG IGCCC					
		GACGGTGAAG GCCCATGTTC					
		CGCCTGGACG					
		CGCGGTGACC					
50	2281	GTAGGTGTGC	GATGTGCCCG	CCCTCAGGCC	GGTGTCCGTG	TACGACGTCG	TGGCGGACGT
	2341	GGTGATCTGG	GCACCGTCGC	GGTGGACGGC	GTAGTOGGTG	GCGCCGTCGA	CGGGTTTCCA
	2401	GGTCAGGCTG	ATGGTGGTGT	CGGTGGCGCC	GGTGGCGGCC	AGGCCGGACG	GAGCGGGCAG
		CGAACCGGGG				CACTGCTGTG	
55	2521 2581					TTCACCTCCA	
23		TOCGTOCGCG					
	2701	GTCCGGCGTC	TGGGACACGC	CGTGCACAGC	GGTCCACTGG	TCGCGCAACT	CGTCGGCGTT
	2761	GOGOGGGGGG	ACGGTGGTGT	CCTTGTCGCC	GTGCCAGATG	GCCACGCGCG	GCCACGGGCC
<i>c</i> . O		CGACCACGAG					
60	2681 5041	CCCGGGGTTC GGCGACGACC	ATGCACAGGT	ACGUGUTGUT	- GAGETEGERG - OGGETEGERG	GCALAGUUGA	CCGACGTCAT
	294I	JURUJAULUU	300000001	SUMMONUGIU	0001170013	J001.501110A	CCC.TCGTCA:

	2221	200100000	2222222		2000000000	222222222	CCT A CCCCC A
		GGCACCGCCG	GCGGACAGCC	CoolGAIG.A			
	3061		GCGGCCATCT			CCCTGGCCCC	
	3121	GOTGOTGTGG	PACCAGTIGA	AGCACCTGTT	CGCGTTGTTC	GACGACGTGG	
	3191	CACGAGCAGG		GGTCCGCGAA	TGAGAGCAGG	CCGGAGTTGT	CGGCGTAGCC
5	3241	0130303110	TGGGTGGAAC	CGTGCAGGGC	GRACACCACC		CGGGCAGGGA
	3300	0300300003	TAGACGTACA	TGTTCAGCCG	3000333770	STGCCGAAGT	CCGCGACCTC
	3361	GGTGAGGTGG	SCCTTGGTCA	SADOSGGGIT	3300%36000	GCCGCGGCGT	GGGCCGTCGG
	3421	cacchaacca		STOCGAGTAC		ACGGCCACGA	GACGGGMGAG
	3481	CACCCCCCGC	CGTCCCGGAC	GOGACAACGA		GGCGAGGAGG	
10	3541			GGRACGGCGG			CGATGTCGTG
. 0	3601		GGAGGGCTCC	CTGACGTCGA		CGCCCCGGTG	
		GGGGGGACAC				ACACCCCGCA	
	3661	TAGGGGTGGT	TCAACCCGCA				
	3701	TGCGCCCGGA		CGCCTTGCGG		CCGGACGCGA	
	3781		GGTAGGGCGT	CATGGTGTCC		GTCGGCCTTG	
15	3841	ACGGACCGGG	CGTCGGCGGA		GCGGGGTGGG	CGGTATGGCG	
	3901	CCAGCCGCGT	GGGGGGGGGG	CGCCCAAGTG		ACCGTGGCCG	GCGGGAGGGC
	3961	CGGACCGGTC	ASTGCAGTCC	CGCGGCCCTG	CGGGACCGCT	COTCCCAGAC	GGGTTCCACC
	.021	GOGGCGAACI	00000100010	TULGGGGGGG	TAGACCATCA	STGTCCGCTC	GAAGGTGATG
	4081	ACGATGACAC	CGTCCTGGTT	GTAGCCGATG	GTGCGCACGC	TGATGATGCC	TACGTCAGGT
20	4141		ACTCCCGGGT	GTTCAGGACC	TOGGACTGOG	AGTAGATGGT	
	4201	AAGACCGGGT		GACCCGGTCC	CAGCCGAGGT	TGGCCATCAC	
	4261	ATGTCGGTGA		GGTGACCAGG		AGGTGGAGTC	
			TGGTGCCCGC			GCAGCGGCGC	
	4321						
2.5	4381		TGAGCCAGGA			TGCGGCCCAG	
25	4441		CGGTGGTGAA		TAGOGGGGGT	GCCAGCCCTC	
	4501	GTGCGGGTGG		CGGGTTCTCA		CGCTCATTCT	
	4561	CGGTCCGCTG		AACCTTCACC		GTGCGGCGCA	
	4621	ACCGTACGTA				TGGTCCTCCG	GCGAGTGTGA
	4681	CCACGCCGAC	CGTGCGCCGC	GCCTGCGGGT	CGTCGAGCGG	CACGGCGACG	GCGTGGTCAC
30	4741	CGGGCCCGGA	CGGGCTGCCG	GTGAGGGGGG	CGACGGCCAC	ACCGAGGCCG	GCGGCGACCA
	4801		CGTGCTCAGC			CCGCGGCACG	AATCCGGCCG
	4861		CCGGTCGGTG			CGGCTCCAGT	GCCACGAACG
		CCTCATCGGC				GGCCAGCCGG	TGTCCGGGTG
		GGACGAGCAG					
35	5041	CMCCMCCCCM	GGTCAGCCCC	7.007.000.000	TOCTOTOCO.	GACGTCGTCG	ACCACGGCGT
33							
						CAGCCGGCGG	
			CACCAGCCAG				
	5221		GATCAGGGCG			GGAGACCTCA	
	5281	GCAGGGCGTG	GGCGCGGAAG	ACCTCGCCGT	ACTTGTTGAG	CCGGAGCCGG	TTCTGGTGCC
40	5341	GGTCGAACAG	CGGCACGCCC	ACTCGTCGCT	CCAGCCGCCG	GATGGCCCTG	GACAGGGTCG
		GCTGGGAGAT					
	5461	TGAACCACTG	CAACTCCCGT	ATCTCCATGC	AGGGACTATA	CGTACCGGGC	ATGGTCCTGG
	5521	CGAGGTTTCG	TCATTTCACA	GCGGCCGGGC	GGCGGCCCAC	AGTGAGTCCT	CACCAACCAG
	5581	GACCCCATGG	GAGGGACCCC	ATGTCCGAGC	CGCATCCTCG	CCCTGAACAG	GAACGCCCCG
45	5641	CCGGGCCCCT	GTCCGGTCTG	CTCGTGGTTT	CTTTGGAGCA	GGCCGTCGCC	GCTCCGTTCG
	5701	CCACCCGCCA	CCTGGCGGAC	creacecec	GTGTCATCAA	GATCGAACGC	CCCGGCAGCG
	5761	GCGACCTCGC	CCCCCCCCTAC	GLCCGCACGG	TOCGTOGCAT	GTCCAGCCAC	TTCGTCTGGC
	5001	TGAACIGGGG	CDACCACACA	CTCCACCTCC	AMONGOGOTO	GCCGGAGGGC	AACCGGCACC
	2051	TGCACGCCTT	CARGGAGAGC	G.CCAGCTCG	TCTTCCACAA	TOTOCOACCC	CCCCCCCCCC
50	2021	1GTACGCC11	T T T T T T T T T T T T T T T T T T T	GCCGWIG.2C	CCCCCCCC	CCLCCCTCTT	CACCACCCC
50	5941	GCCGCCTGGC	ATCGGCCACC	AGGICCICGC	GCGGAGCCAC	CINCCCCENC	CRCCTGCGGA
	6001	CATATCCGGC	TACGGCAGTA	CCGGCTGCTA	COUCUSACOS	CAAGGCGIAC	GACCICCIGG
	6061	TCCAGTGCGA	AGCGGGGCTG	GTCTCCATCA	CCGGCACCCC	CGAGACCCCG	TCCAAGGTGG
		GCCTGTCCAT					
	6181	TGCTGAAGCG	GGCCCGCACC	GGCCGGGGCT	CGCAGTTGGA	GGTCTCGATG	CTCGAAGCCC
55	6241	TCGGTGAATG	GATGGGATAC	GCCGAGTACT	ACACGCGCTA	CGGCGGCACC	GCTCCGGCCC
	6301	GCGCCGGCGC	CAGCCACGCG	ACGATCGCCC	CCTACGGCCC	GTTCACCACG	CGCGACGGGC
	6361	AGACGATCAA	TCTCGGGCTC	CAGAACGAGC	GGGAGTGGGC	TTCCTTCTGC	GGTGTCGTGC
	6421	TACAACGCCC	CGGTCTCTGC	GACGACCCGC	GCTTTTCCGG	CAACGCCGAC	CGGGTGGCGC
	6181	ACCGCACCGA	GCTCGACGCC	OTGGTGAGCG	AGGTGACGGG	CACGCTCACC	GGCGAGGAAC
60	6211	TGGTGGCGCG	COTTO A COTTO	GOGTOGIONES	COTACGCACG	CCAGCGCACC	GTGCGGGAGT
00	2001	TCAGCGAACA	CCCCCA A CEC	- CCACACCCAC	CACCCTCCCC	TCCGTTCGAC	AGCCCGGTCG
	0001	T CAGCGAACA	CCCCCAACIG	CCCCCCCCCCC	CCTTCCACCC	CONCORRECCE	CCCCCCCCCCCCC
	0001	GTGCGCTGGA	GGGCCTGATC	CUCCUGGTCA	T CT CCACGO	CCCCTCCCTC	0000000000
	6/21	GCCGGGTCCC	GGAGCTGGGC	GAGCATACCG	AGICCGTCCT		
	6781	ACAGCGCCGA	CCGCGAAGAG	GUUGGCCAT'G	CUGAATGAAC	rcaccggagT	CCTGATCCTG

	6841	GCCGCCGTGT	TOCTGOTOGO	CGGCGTACGG	GGGCTGAACA	TGGGCCTGCT	CGCGCTGGTC
	6901	GCCACCTTTC	Tecrosect	CCTCGCACTC	GACCGAACGC	CGGACGAGGT	SCTGGCGGGT
	5951	TTCCCCGCGA	GCATGTTCCT	3373275372	GCCGTCACGT	TECTETTEGG	GATCGCCCGC
	~ 021	GTCAACGGCA	CGGTGGACTG	GOTGGTACGT	gresessarse	SGGCGGTGGG	GGCCCGGGTG
5	7.93	COMSCORTEC	COTESETSCT	0770000070	30330A3730	TOTGCGCGAC	AGGCGCGGCC
	77 2 3 2	1030000003	20075000AT	03733300000	ATCABABING	CGTTCGCCGT	CAGGCACCGC
		ATOGATOOSO	TOTACGCCGG	ADTGATGGCC	3T3AA33333	COGCAGCOGG	CAGTTTCGCC
	7261	JCCTCCGGGA	TCCTGGGCGG	CATOGTOGAG	1099090100	AGAAGAACCA	TOTGCCCGTC
10	7321	ASCGGGGGGG	racrotrosc	AGGCAGGTTC	GOOTTONACO	TGGCGGTCGC	CGCGGTUTCA
10	7391	TOGGTCG	TOGGGCGCAG	GOGGO COS	Comunication	TGGACGAGGA	CACCGATCCC
	7441	ACGGAAGGGG	ACCCGGCTTC	0000000000	3003AA0A00	TGATGACGCT CCGGCTTCCT	GACCGCGATG GGCCCTCACC
	7501	TIGGCGGCTGG	TGCTGGGAAC	GOTOTTOCOS	OGCACCTCCC	AGCAGGCCAC	CAAGGAGATC
	7561	*********	TGCTGGCGCT	GGTATGCGGG	*********	ACGTCGCCCT	GCTCCAGGAG
15	7621 7681	DIBGGGGATIG	TIGGTGCTGCT	GGGGAAGATG	ATCGCGGCGA	TOGGCACOCO	GCTGCTGGCC
15	7741	GCCCTGGTGA	TOTGOTACST	GGGGGGTGTC	30000000000	TOGOCTOGAC	CACCGGGATC
	7801	CTCGGTGCCC	TGATGCCGCT	GTCCGAGCCG	7700705507	CCGGTGCCAT	CGGGACGACC
	7861	GGCATGGTGA	TGGCCCTGGC	ggccadaaca	ACCGTGGTGG	ACCCGACTCC	CTTCTCCACC
	7921	AATGGTGCTC	TGGTGGTGGC	CARCGCTCCC	GAGCGGCTGC	GGCCCGGCGT	GTACCAGGGG
20	7981	TIGCTGTGGT	GGGGCGCCGG	GGTGTGCGCA	checonocce	CGGCCGCCTG	GGCGGCCTTC
	8041	GTGGTGGCGT	GAGCGCAGCG	GAGCGGGAAT	CCCCTGGAGC	CCGTTTCCCG	TGCTGTGTCG
	8101	CTGACGTAGC	GTCAAGTCCA	CGTGCCGGGC	GGGCAGTAGG	CCTAGCATGT	CGGGCATGGC
	8161	TAATCAGATA	ACCCTGTCCG	ACACGCTGCT	CGCTTACGTA	CGGAAGGTGT	CCCTGCGCGA
	8221	TGACGAGGTO	CTGAGCCGGC	T000003000A	GACGGCCCAG	CTGCCGGGGG	GTGGCGTACT
25	6281	SCCGGTGCAG	GCCGASGAGG	GACAGTTCCT	OGNOTTOOMG	GTGCGGTTGA	CCGGCGCGCG
	5341	PORGGTGCTG	GAGAT CGGGA	COTADACCOS	OTACASSASS	crereceres	CCCGCGGATT
	8401	330GCCCGGG	SSCCGTGTGG	TGACGTGCGA	TOTOATOONS	AAGTGGCCCG	AGGTGGGCGA
	8461	GUGGTACTGG	GAGGAGGCCG	GGGTTGCCGA	COGGATCGAC	GTCCGGATCS	GCGACGUCCG
3.0	8521		ACCGGGCTGC	TOGACGAGGC	GGGCGCGGGG	CCGGAGTCGT	TCGACATGGT
30	8581			CCGGCTACCC	CGCCTACTAC	GAGGGGGGG	TGCCGCTGGT
	8641	ACGCCGCGGC		TCGTCGACAA	CACGOTGITC	TTCGGCCGGG	TGGCCGACGA TGCGCGACGA
	8701		GACCCGGACA		ACGCGAACGC GGCCGACGGC	AACGCGGCAC GTCACCCTGC	TGCGGAAACG
	8761 8821	CGACCGGGTG		TGCTGACGAC GGCGGTCAGC	GTCAGCGTCG	TOGGCGCGGG	CCTCGCGGAG
35	8881	GTGACCGGGG GGCTCCAGAT	GCAGGCGTTC	GACGCCGGCG	GCGGAAGCGC	CCGCCACCTC	GGACACGCAG
23	6941	GGGCAGTCGG	AGTCCGCGAA	GCCCGCGAAC	CGGTAGGCGA	TCTCCATCAT	GCGGTTGCGG
	9001	TOGGTACGCO	GGAAGTCCGC	CACCAGGTGC	g0000000000	GGGGGCCCTG	GTCCGTGAGC
	9061	CAGTTCAGGA	TOSTOGOACO	GGCACCGAAC	GACACGACCC	GGCAGGACGT	GGCGAGCAGT
	9121	TTCAGGTGCC	ACGTCGACGG	offorford	AGCAGGATGA	TGCCGACGGC	GCCGTGCGGG
40	9181	CCGAAGCGGT	CGCCCATGGT	GACGACGAGG	ACCTCATGGG	CGGGATCGGT	GAGCACGCGC
	9241	GCAGGTCGGC	GTCGGAGTAG	TGCACGCCGG	TOGOGETTOAT	CTGGCTGGTC	CGCAGCGTCA
	9301	GTTCCTCGAC				GATCGTCATG	
						CCGGGCCTGG	
	9421	AACCCGCCTG	GTACATCAGG	CGGCGCGGAC	GCGAGTCGAC	CGTGGACACC	GGCGGGCTGA
45	9481	ACTCCGGCAG	CGACAGGAGC	GTGGTCGCCT	GCTCGGCCGG	GTAGCACCGC	ACCTCGGGCA
	9541	GGTGGAACGC	CACCTOGGCA	CGCTCGGCGG	GCTGGTC3TC	GATGAACGCG	ATCGTGGTCG
	9601	GTGCGAAGTT	CAGCTCCGTG	GOGATOTOGO	GGACGGACTG	CGACTTCGGC	CCCCATUCGA
						CAGACGCTCC	
50	9721	CGTGGTCGTT	CTTGCTCGCC	ACCGCCTGGA	COMPONENCE	GTCGTCGAGC CAGCACGGTG	CCCCCCACA
30	9781	COTOGCGGAT	CICGICGGIG	AGGACCACCI	mononameum	CATGGCTGTC	COCCGCCACA
	9541	AGG.GIIG.C	CAGGICCCAG	ACCASACAU.	ACCCCCCCC	TCTCGCTGCT	GCCCTCGATG
						CGTGTCCCTC	
						CGGCTCGTTC	
55	1002	TGCTTGGCCA	GGATCGTCGC	GGGCACCATC	TOGGGCGAGO	CCTCGTCCCA	GTGGTCGCTG
	10141	GCGTACTCGC	ACACGCGGGC	CGCGATCTGC	TCCGCGGTCC	ACAGGTCGGC	GATGTGCCCG
	10201	GCGACGAGTT	GGTGGTCGCC	GAGCGGCCGG	CCGAACTGCT	CCCGGGTCCG	GGCGTGGGCC
	10261	ACCGCGGCGG	TGCGGCAGGC	CCGCAGGATC	CCGACGCAGC	CCCAGGCGAC	CGACTTGCGC
	10321	CCGTAGGCGA	GTGACGCCGC	GACCAGCATC	GGCAGTGACG	CGCCGGAGCC	GGCCAGGACC
60	10381	GCGCCGGCCG	GCACACGCAC	CTGGTCCAGG	TGCAGATCGG	CGTGGCCGGC	GGCGCGGCAG
	10441	COGGACGGCT	TCGGGACGCG	CTCGACGCGT	ACGCCGGGGG	TGTCGGCGGG	CACGACCACC
	10501	ACCGCACCGG	AACCATCCTC	CTGGAGACCG	AAGACGACCA	GGTGGTCCGC	GTAGGCGGCG
	10561	GCAGTCGTCC	AGACCTTGTG	GCCGTCGACG	ACAGCGGTGT	CCCCGTCGAG	CCGAACCCGC
	10621	GTCCGCATCG	CCGACAGATC	GCTGCCCGCC	TGCCGCTCAC	TGAAGCCGAC	GGCCGCGAGT

	10681	"TCCCGGCTGG	mengercerr	CAGGAAGGTC	GCCCGCTGAC	CGGCGTCGCC	GAGCCGCTGC
	-0-4-	recense	COCCUTACC	orgogaegre	ATGACACTGC	GCAGCGAACT	GCAGAGGCTG
	10801	COGACGTGTG				GTCCCAGACC	
-		GOOGCCACTT		CAGGCCGTCG		GGACGAGCAG	
5	10901	AGTTCGCCGG		CTCGGGGGGC		CAAGGTCGGT	CAGCAGCGCG
	10981	TOROGOTORG		CCCGASCCG		GCGACCATGG	
		ACGGAAGITO	303A30T33A	G000033300	GGCGATCGTG	ACGTCGAACG	TCTTCTCCAG
		STACASSASS		CGARCAGOGA	COTGRESCOS	CCCTCCGCGA	ACAGGTCGCG
		STOCASSSC			CTTGAGGAAC		
10		GGGGTCGTCC				CGTATTCGTA	
10							
						GCAGCAGGTC	
		CTGCGCTCGT				CGTCGACGAG	
	11461	COGATCAGGT	CCGCGGTGCG	CAGOGGCCCG	GTCGGATGGC	CGAGGCACCC	CGTCATGAGC
	11461	GOGTOGACGT	CCTCGACGGA	CSCGGTGCCC	TCCTGCACGA	TCCGCGCCGC	GTCGTTGATC
15	11521	ATCGGGGTGGA	GCAGCCGGCT	COTGACGAAG	CCGGGGGGGGT	CCCGGACGAC	GATCGGCTTG
	11581	CGCCGCAGCG		amendedaged	GCGGCCATGG	COTTCTCACC	GGTCCGGGGT
	11641	COGCGGATCA		CGGGATCAGG	TACGACGGGT	TCATGAAGTG	
	11701				AGTTCGTCAA	CCGGGATCGA	
		AGGTCCTCGG					
• •	11761	STGATGACCS		CGCCGCTGCC	GAGACCGTGG	CGAGTACCTC	
20	11821	POSSCGTCCT	CGACGACGGC	CTCGATCACC	GCGGTGGCCG	TACCGATCGC	GGGCAGCGCG
	11991	GACGTGGCC3	TOOGOAGCAC	ACCGGGGGTCG	GCCTCGGCGG	GCCCGGCCAC	GAGTTGTGCC
	11341	GTCCGCAGTT	CGGTGGCGAT	CCGCGCCCGC	GCCGCCGTAA	GGATOTCCTC	GGACGTGTCG
		ACGAGTGTCA		GTGGCGCAGC		TGATGCCGGT	
		CCCGCGCGCA				CTCCCTCCGG	
25		GCAGCGAGTA				GATCGCGTCC	
20							
		GGCCGAGTTC		COGAGTIGCA		GATGTGGTCG	
	12241	TGCCCGTCGA		CTCAGGCTGT		CGCCGCGGTG	
	12301	CGCACAGGGG		GGGCCGAGCT		CAGTTGCTGG	
	12361	CGGCGCGGGC	CTGCCCCGGA	TGGTCGACGC	AGATGAACGC	GTCGTCGAGC	AGGGTCTTCG
30	12421	GCAGTTCGGT	CTTGCCCGGC	TOGTOGGOGO	CGATGGCGTT	CACATGCAGG	TGCGGCAGCC
-		GCGGCTCGGC					
		CATCCGCGGC				GACCGGCAGT	
		CGATGCGGTC				GTCGCTGACC	
a -		CGATGGGCAG					
35		TCAGCGTGAG				GCTCGCGACG	GCGGCGACCG
	12781	CGCCGGTCCG	CATCGCGGTG	ATCACGCCTG	CGTCGGCGAG	GGCGGTCAGA	CTGCCGCTGT
	12841	CGTCGTCGAG	GOGOGACATO	GTGCCGACGA	TCGTCGGCAG	CCGGAAGCGC	GGATAGTTGT
		GCGGACTGTA		TTCATGGTCA		GGGGACCCGG	TACGGCATGA
		ACTCGATGAC		TOGOCGOCGO		GGTACGCGGC	
40						CTCGCTGATC	
-+ 0							
						GTCACGTTGG	
						CTTGGTGGTG	
						AAAATCTCGT	
						TCCCGCCGCC	
45	13321	CAGGGCGTCC	AGCCGGGTTC	CGATCGCGTC	CGCCTGGCGG	GCGCCCGGGT	CGACACCGGC
	13361	AACGAGTGCT	"CCAGCCGGT	CSAGCTGCGC	GAGCACCACG	GTCACCGGGT	CGTCCGGGGA
						GGGTAGTCGA	
						CGCAGCTGCA	
<i>5</i> 0						ATGTCCTCCG	
50	13621	GCCCAGGACG	GCCGCTGCCT	TCTGCCGGAC	GAGGGCGAGC	AGGTCGGTGG	GGCGTTCCTG
						CCACGCAGCA	
	13741	CGGCGGCAGG	TCGCCCGCCA	CGGCGACGAC	ACTGCCCGTT	CCGGTGTGGA	CGGCGGCGTC
	13801	GTACATGCGC	ATGCCCTGTT	CGGCGGTGAG	CGCGCTCGCC	CCACCCTTGC	GCATACGGCG
						TGGTCCCACA	
55	13001	C000100333	COMCCCACCC	CTTCTCCACC	COGGGGGGTTCS	GCGAGCGCGT	CGAGGAACGC
23	10001	GATCGACAGC	CCIGGOAGCC	CITGIGGAGG	CCGGIGIICG	CONCOCCC	CCCACCAGTA
	13981	G110G00300	GCGTAGTTGC	CUTUACCUG	GG 1GCCCAGC	ACACCGGCCG	N C C C C C C C C C C C C C C C C C C C
	14041	GACGACGAAT	GCGGCGAGGT	CGGTGTCGCG	GGTGAGCCGG	TGCAGGTGCC	AGGCGGCGTC
	14101	GGCCTTGGGT	TTGAGGACGG	TGTCGATGCG	GTCGGGGGTG	AGGTTGTCGA	GCAGGGCGTC
	14161	GTCGAGGGTT	CCGGCGGTGT	GGAAGACGGC	GGTGAGGGGT	TGAGGGATGT	GGGCGAGGGT
60	14221	GGTGGCGAGT	TGGTGGGGGT	CGCCGACGTC	GCAGGGGAGG	TGGGTGCCGG	GGGTGGTGTC
-	14281	GGGGGGTGGG	GTGCGGGAGA	GGAGGTAGGT	GTGGGGGTGG	TTCAGGTGGC	GGGCGAGGAT
	1/2/1	000000000	STOCCOUNTS:	CGCCGGTGAT	GACGACGGCC	CCCTCGGGGT	CCAGCGGCCG
	14341	0000000000	TACTOCAMEN		CTCCCCCCCC	CTCATGGTCG	CCAGCGCCTC
	14401	CGGGACCGTG	AGGAUGATOT	100000101G	CICGCCGCGC	CICHIGGICG	CCCCCCCC
	14461	GUGGACCTGC	CGCATGTCGT	GUACUSTUAC	CGGCAGCGG	TGCAGCACAC	CGCGCGCGAA

	14521	0100000100	* 00ma00000	mojimamaanm	GAGCCGGTCG	222222222	CC N TIC N CCTC
	14581				OGMOTT0000		
	14641	GAAGGGT0GC	TGGACGGCGT AGGCCGACGG	- 200000000	03.01.0000	GTGAGCGAGT	ACCGGCCACC TGAGCACGAC
		3703A03GGC	•	6380GAA060	GGTGCTGCGG	SAATOGGCCA	GATGCGCTCC
5	14761	GICCAGGICC	ACCAGATEGO	G017030G30	Campanacaa	GOGTACACCT	CCGCGCCCAG
,	14521	3100033100 3100033300		00000001100	GACACCGCCG	GIGGCCGCGT	GGATCAGGAC
	14581	(2) (2) (2) (2) (2) (2) (2) (2)	GGGGGGGAGGG	03303AG3T 1	GACCAGGCCG	TACCACGCGG	TOGOGAACGO
	14341		GACGCCGCCT	00000011000	2210000000	0001000000	CGAGCATCCG
	15001	gmormosone	ATGACCGTGG	alboorunid. Alboorunida		AGGCCGAAGA	CGCGGTCGCC
10	15081	-2-33:33-00 -55555555555	200GAGACGT	03303000031	OTOCAGGACS	ATRICCOGOGO	CSTGGCGGCC
10	15121	GAGGAGGGGGG	TGACCGGGGT	AGGTGCCGAG	COCGATCAGO	ACATCGCGGA	AGTTGAGGCC
	15161	CGCCGCACGC	ACACCGATCC	001000000	0303000433	GGGGGGGGG	GCTCCGCCGA
	15241	3TCGG00000G	GTGAGGCCGT	CGAGGGTGCC	CGTCCGCGCCC	GGCCGGATCA	GCCACGTGTC
	15301	COTGTCCGGC	ACGGTGAGCG	GCTCCGGCAC	CCGGGTGAGG	CGGGCCGCCT	CGAACCGGCC
15	15361	000000000000	CGCAGACGCG	GCTCGCCGAG		ATGCGCTGCT	GCTCGGGGGC
15	15421	GAGCGTGACG	COGGACTOGG	TOTOGACGIG	GACGNACCGG	COGGGCTGCT	CGGCCTGGGC
	15481	0300010AGC	AGTECGGCCG	0030G0CG3T	GGCGAGCCCC	GOGGTGGTGT	GCACGAGCAG
	15541	ATOCCCCCCC	GAGCCGGTCA	GGGCGGTCI	CAGCCGGGTG	STGAGCGCAC	GCGTCTCGGC
	1560:	CACCGGGTCG	TOGCCATCAG	CGGCAGGL.FA	CGTGATGACG	TOCACGTCGG	TCGCGGGGAC
20	15661	ATCCCTGGGT	GCGGCGACCT	CGATCCAGGT	GAGACGCATC	AGGCCGGTGC	CGACGGGTGG
	15721	GGACAGCGGG		CCGTCCGGAT		AGTTGGCCGG	CGGAGTCGGC
	15781	GACGCGCAGA	CTCAGCTCGT	CGCCGTCACG	AGTGATCACS	GCTCGGAGCA	TGGCCGAGCC
	15841		AACCGGGCCC	CCTTCCAGGC	GAACGGCAGA	COCGCAGOGO	TGTCGTCCGG
	15901	CGTGGTGAGG		GCAGGGCCGC	GTCGAGCAGC	SCCGGATGCA	
25	15961	andescares	GCGGCCTGCT	CGTCGGGCAG	CECCACCICS	GCATACACGG	TGTCACCATC
	16021	ACGCCAGGCA		CCTGGAACGC	CGACCCGTAT		CATCCCGCAG
	16081		AACCCCGAGA	CGTCGACGGC			ACTGCGAGAA
	16141	COGCTCCACA	CCGACAACAC	COSCESTETC	GGGGGTGTCG	GGGGTCAGGG	TGCCGCTGGC
	16201		CAGCIGCCCG	TGCCCTCGGT	ACGCGCGTGG	ACGGTCACCG	GCCGCCGTCC
30	16261	GGCCTCATCA	GCCCCTTCCA	CGGTCACCGA	CACATCCACC	GCTGCGGTCA	CCGGCACCAC
	16321	AAGGGGGGAT	TCGATGACCA	GCTCGTCCAC	TATCCCGCAA	CCGGTCTCGT	CACCGGCCCG
	16381	GATGACCAGC	TCCACAAACG	CCGTACCCGG	CAGCAGGACC	GTGCCCCGCA	CCGCGTGATC
	16441	AGCCAGCCAG	GGGTGAGTGC	GCAATGAGAT	CCGGCCAGTG	AGAACAACAC	CACCATCGTC
	16501	GGCGGGCAGC	GCTGTGACAG	CGGCCAGCAT	CGGATGCGCC	GCACCCGTCA	ACCCCGCCGC
35	16561	CGACAGATCG	GTGGCACCGG	CCGCCTCCAG	CCAGTACCGC	CTGTGCTCGA	ACGCGTACGT
	16621	GGGCAGATCC	AGCAGCCGTC	CCGGCACCGG	TTCGACCACC	STGTCCCAGT	CCACTGCCGT
	16681	GCCCASGGTC		CCAACGCCGT	CAGCCACCGC		CGTCACCGGT
	16741	CCGCAACGAC			CATCGCCGGC		GATGGGCACT
	16801	GCACTCCACG		CATCCAGCTC	CGCCACCGCC	GCGTCCAACG	
40				ACCCCTCATC	CACCGGCTCC		
	16921		CACGCCACCG				
					GTGGGAGGCG		
	17041	CACCCGCACG	CCTTCGGCCT	CATACCGCGC	CACCACCTCC	TCCACCGCCG	ACGGGTCCCC
4.5	17101	CGCCACCACC	GTCGAAGCCG	GGCCGTTACG	CGCCGCGATC	CACACACCCT	CGACCAGACC
45					CATCGCTCCC		
					GGCGGCGTCC		
					GTGTCCGATC		
					CGCGACCGCC		
50					CAACATCTCC		
50					ACGCGCGGCG		
					GCCCTGGCCG GGCATCGCCC		
					CGCGACCGCG		
					CCCCCGCAGA		
55					CGACTCCCCA		
ل ل					CACCCCGAAC		
					CTCGGTGAGC		
						TTCGGCGCGA	
					GACACCCCC		
60					AACCTCACGC		
00					CGTCGTCCCC		
	18121	CTCCACITCC	#000000000 #0000000000	GTCCGGCGTT	CACCAACGCC	TGCTGGATGA	CACGCTGCTG
					GGAGGCACCG		
					GCGCTCGGCG		
	10001	555511561166	200.10.41000				. =

	1000	AAGAACGCCG	0000000000	2007 0200 cm	20000000000	000000000	1 222220001
		GCGGCCGTCG					
		CATGACGGTG					
		GGCCTGGTGC					
5		CTGGAGCCCA	TAGAAGTAGG			CTGGGGTGCA	
	18661	GCCGAACCCG	mantananas			GAGAAGGCGC	
	18721	GCCGGTGTCG					
	18781	TSICSTITES	ACCACACACA	compoundable	STOCATECOS	COTOCOTOLO	CSSSSSTEAM
	18841	GCCGAAGAAC				AAGCCGCCGC	
10		CGATCCGCCG				GCCGGGAAGC	
• -	16961	STOGGGGGGA	CTGTCCACCA	TGCGCCACAG	spegressa	SAGGTGACGC	
		TOGGCAGGCC	ACCCCCCA	TOGGGGGGGG	TTCGT02.033	arcacaaaa	CTSTGGGAAC
	19081	AGCGACCGGT	GCGGCACCAC	CGACCAGAGO	CTGGTCCAAC	CGCGACGCGA	TGGCCCGCGG
	19141						
15		GTTCCGCAGT					
		GGACACGTCC					
		CAGCAGCGCG					
		GGCGGTGGCC					
		STGCGCGGTG				STGCCGGTTC	
20	19501	TTCCAGCAGG	CGCATGCCCA	CACCGGCCGA	CATGGGGGCGG	AAACCGCCGC	GGCGGACACG
	19561	GGTGCGGTTG	GTGCCGCTCA	TGCTGCCGGT	GAGTCCCCTC	TOATOGGCCC	AGAGGCCCCA
	19621	GGCCAGCGAC	AGCGCGGGCA	GTCCTTCGGC	ATGGCCCAGC	GTCGCGAGTC	CGTCGAGGAA
	19681	CCCGTTCGCC	GCCGAGTAGT	TGCCCTGGCC	GCGGCCGCCC	ATGATGCCCG	CBACGGACGA
	19741	GTAGAGGACG	AACGAGCGCA	GGTCCGCGTC	CCGGGGTCAGC	TOGTGCAGGT	GUCAGGCGCC
25		GTCGGCTTTG					
		GTCGTCGAGC					
		CGCGGCGGCG					
		CGCCGGCGGT					
• •		ATGCCGGGCG					
30		CGGGTCGAGC					
		GTACCGGCCG					
		CTCGATGGGG					
		GGCGGACCGG					
25		GAGGGTGGTC					
35		CTCGGTGAGC					
		GATGTGGACC GTACAAGGAG					
		CGCGGCGACG					
	20641					GTCGTGTGGA	
40		GCTCCACGAG					
40		GACGTGCAAG					
		CTGTTCCCCG					
		CAGTCCCTGG					
		GCTCACGTCG					
45		GCTTCCGGCC					
		ACGCGCGTGG					
		CACATCCACC					
		CACCCCGCAA					
		CAGCAGAACC					
50		CCGGCCAGTG					
		CATCGGATGC					
	21421	CAGCCAGTAC	CGCCTGTGCT	CGAACGCGTA	GGTGGGCAGA	TCGAGCAGCC	GTCCCGGCAC
	21481	CGGTTCGACC	ACCGTGTCCC	AGTCCACTGC	CGTGCCCAGG	GTCCACGCCT	GCGCCAACGC
		CGTCAGCCAC					
55		TTCCATCGCC					
		CTCCGCCACC					
	21721	ATCCACCGGC	TCGGTCACCC	AGGCGCTGTC	CACCGTGGAC	CACCAGGCCA	CCGACCCGGT
		CCCGCCGGAA					
		CGTGTGGGAG					
60		CGTCACCACT					
		ACGCGCCGCG					
	22021	AGCCATCGCC	CCCCGCCCGG	CCAGCCGCCC	GGCGATCACC	TGGCTGCGCA	AGGCCACCAC
	22081	GCGGGCGGCG	TCCTCAAGGC	TGAGGGCTCC	GGCCACACAC	mcccacacaca	1C1CGCCCTG
	22141	GGAGTGTCCG	ACCACCGCGT	CCGGCACGAC	CCCATGCGCC	TGCCACAGCG	COGCCAGGCT

				22222000000	0100000000	222022222	ON WOOCCOOC
	22201	CACCGCGACC	GCCCAGCTGG	CCGGCTGGAC		CGCTCCGCCA	
	121161	COCCAACATO	TOCOGOAGAT	00000000000	GTUCGGCAAC	•	CACACTCCTC
		CATACGAGCC	BOSAACACOG	UABANUACGO	CALCAACICC		CCACCCACTG
-	02381	ACCACCCTGC	CCGGGAAAGA	Connunt Cacce	AUGUSSULUM	TCCACCGCCA	
5	22441	COGGGGCATEG	CCCAACAACA	cogcacgana	AUGUMMUMUM		GCACCAACCC
	22501	chacacaca	GOGGOCACAT	UUACACCAUU	COCCECAN		GCCGCTCCAC
	23581	STGSCSCGSS	AGACTCACCT	CACTCCGAGG	CGACACCCCCC	HACGGCACCA	
	22621	AGCOGACTCC	COACGCGACG	GCCCGGGAAC	ACCOTCAAGG	= =	CGTTCGTACC
• •	22661	GCTCACCCCG	AAAGCGGAGA	CACCGGGCCCG	GCGCGGACGT		GCCACGCCCG
10	22741	caccrossis	AGGAGTTGCA	cogdatacha	GGTCCAGTCC		ACGGCTCGTC
	12801	CACATGCAGC	GTCTTCGGCG	CGATGICATA	COGCATCCCC		TGATGACACC
	22861	GGCGACACCC	GONGCOGCCT	GCGCATGACC	GATGTTOGAC		CCAGCAGCAG
	22921	DEGRACOTOR	CGCTCCTGCC	CGTACGTCGC	CAGAATOGOG		TGGGATCGCC
	22981	CAGCGTCGTC	CCCGTCCCGT	GCGCCTCCAC	CACOTCCACG		CGAGCCUCGC
15	23041	CTTGTGGRGG	GCCTGGCGGA	TGACGCGCTG	CTGGGAGGGG	COGTTGGGTG	CGGAGATGCC
	23101	STTCGAGGCG	CCGTCCTGGT	TGACGGCGGA	GGAGCGGACD	ACCGCGAGGA	CGGTGTGTCC
	23161	9779777773	GCGTCGGAGA	GCTTTTCGAC	SACGAGGACA	000000007	CGGCGAAACC
	23221	ggrdc	GCCGCGTCAG	CGMACGCCTT	33233	TOOGGGGGGGA	CGCCGCCC15
	23281	COGGGGAGAAC	TOCACGAAGG	TOTOTOGTGA	TGCCATCACT	GTGACACCAC	CGACCAGCGC
20	23341	CAGCGAGCAC	TOCCOGGTCC	GCAGCGCCTG	CCCGGCCTGG	TGCAGCGCGA	CCAGCGACGA
	23401	TGAACACGCC	GTGTCGACCG	TGACCGCCGG	ACCOTOCATE	CCGAAGAAGT	ACGACAGCCG
	23461	TCCCGCGAGC	ACCGCGGGCT	GTGTGCTGTA	GGCGCCGAAT	COGCCCAGGT	CCGCGCCCGT
	23521	GCCGTAGCCG	TAGTAGAAGC	CGCCGACGAA	GACGCCGGTG	TOGOTGOOGO	GCAGGGTGTC
	23581	CGGCACGATG	CCGGCGTGTT	CGAGCGCCTC	CCAGGCGATT	TOGAGGAGGA	TCCGCTGCTG
25	23641	CGGGTCGAGT	GCGGTGGCCT	CGCGCGGACT	GATGCCGAAG	AACGCGGCAT	CGAAGTCGGC
	23701		AGTGCGCCGG	CCCGCCCGGT	GGCGGACTCG	GOGGOGGGT	GCAGCGCGGC
	23761		CCGCGGTCGG	TGGGGAAGTC	GOOGATOGOS	TOGOGGGCGGT	CCGCGACGAG
	23821	CTGCCACAGC	TCTTCCGGTG	AGGTGACGCC	GCCCGGCAGT	CGGCAGGCCA	TGCCGACGAC
	23881	GGCGAGCGGC	TCGTTCGCCG	CGGCGCGCAG	CGCGGTGTTC	TCCCGGCGGA	GCTGCGCGTT
30	23941		GACGTCCGCA	GCGCCTCGAT	CAGGTCGTTC	TOGGCCATCG	CCTCATCCCT
20	24001	TCAGCACGTG	CGCGATGAGC	GCGTCTGCGT	CCATGTCGTC	GAACAGTTCG	TCGTCCGGCT
	24061	CCGCGGTCGT		GGTGCCTGTG	CCGGTGGTTC	ACCSCCGTCC	GGGGTCCCGT
	24131	TGTCGTCCGG		ACGTCCGGGG	CCAGGAGGGT	CAGCAGATGA	CGGGTGAGCG
	24181	CGCCGGCGGC	GGGATAGTCG		TGGCCGGCAG	CGGAATGCCG	AGGGCCTCGG
35	24241	AGACCCGGTT			GCGAGTCGAC	CCCGAGGTCC	TTGAACGCCG
00	24301	TGGTGGCCGT	GACCGCCGCC	GCGTCGGTGT	GGCCCAGCAG		GTGTCGCGGA
	24361	CGACGCCGAG	CAGCACCTGT	TOCOGTTOCT	TGTGGGGCAG	GTCCGGCAGG	CGTTCCAGCA
	24421	GGGAGCCGCC	GTCGGTCGCG	GAGCGCCGGG	TGGGGCGCTG	GATCGGTCGC	CACAGCGGTG
	24481	ACGGGTCGCC	GGGCCCGGGT	GGGGCGGTCG	CCACGACCAC	GGCTTCCCCG	GTGGCGCACG
40	24541				CCGCGGCGGT	GAACGCCACG	GCCGGCAGGC
40	24247	- COGCGTCGAC	0.000100010	accusacce"	GGAGCGGTCC	GGCCGCCTCG	
	24001	CGGCGAGAAC	CANCCCCCCCC	TERMEDICAL	CGCGGGGCAG	SCGGTGCAGT	TOCOAGGOOG
	24001	ACTOGGOGGT	CARCUCCOCC	TOGLOGACOG	COSTOLOGG	GGTTTCCGGC	ACTGTGCCCG
	24/41	GCTCGTACCG	CAMCACAMACC	COSCOTOTO	CCCCSAGGTG	TOCGGOGAGT	TCCTCCGAAC
45	24/01	CGCCCGCGAG	CACCACCCTC	TOCCCCCTT.CC	19900000000	COMEGREGE	GCGGCGGGGA
43	24001	CGAGGCGGGG	CCCTTCCACC	CCCCCCTCGG	CCAGGCGCAG	STECGETTCG	TCGAGGCGGG
	24301	AGAGGCGGC	CCCCTCGAGG	CCCCTCACCC	TOTOGGGGGGG	CTCCACGAGC	ACGAGCCGGC
	24301	CCGGTTCCGC	CCTCTCCACC	NOTICE COLOR	0630200660	GACGGGCCCG	GCCTCGGCGG
	25021	ACACCACCAG	GGTGTCGAGC	AG I GCCCCCA	COMPONENT	TACAGTACA	ACCTCGTCGG
50	25081	GACCGGATAC	CGTGGCGCCG	##C#CCTCC	COCTOCCO	amegeegagg	TCGGTGTACC
50	25141	GGCGGGCCGT	CGGGACGACG	ATGACGTCGG	- GCG199C91C	SETCOLORGE	CGCCGGAACA
	25201	GGCGGGCCGTC GCCGCACGTC		COCCTCCTCC	COCCONTOCCO	GOTGLTGAGG	GAGCCG.TCT
	25261	GECGCACGTC	CCCGTCCGGG	TCCCGICGICG	CGGGGGGGGGG	COCCCCCCCC	TOCCOCTOC
	25321	GAGCCACCGG CGTGGACGAA	CCGTCCCAGT	1CG1CGCGA	CCCCCCCCCC	CTCCACACCC	ACCCCCCTGA
e e	25381	CGTGGACGAA	GGTGACGCGC	AGTITUGIGG	CGCCGCTGG.	CCCCATCCTC	CCCCCTTGCA
55	25441	ACGCGAACGG	CAACCGTACC	CCCGCGTTCT		CCCCAIGCIG	CCCGCTTGCA
	25501	GCGCGGTGAC	GAGCAGCGCC	GGGTGCAGTG	TGTAGUGGG	CCACCCCCCC	CACATOGGGGG
	25561	CGTCGAGGGC	GACTTCGGCG	CAGACGGTGT	CIUUGIGGI		GWCWI GCCGC
	25621	. GGAACTCGGG	GCCGAACTCG	TATCCCGCGT	CGTCGAGTCG	CIGGTAGAAG	TOCOTOTTO
	25681	CGACCGGTTC	CGCGTGCTCG	GGCGGCCAGG	GUCUUGGUGT	GB I GGCCGGT	4.CGG1GG1GG
60	25741	CGATGCCGGC	GAAGCCGGAG	GCGTGGCGGG	TCCATGTCCG	Gruguaga	COCROCROCR
	25801	L GGACGCGCAC	GGCACGGCGT	CCGGTGTCGT	CGGGCGCGGC	GACGGTCACG	CGCACCTGGA
	25861	CGGCGCCGGT	GGCGGGCAGG	ACCAGCGGTG	TOTOGACGAC	CAGTTCGTCG	AGUAGGTUGU
	25921	AGCCTGCCTC	GTCGGCGCCG	CGTCCGGCCA	ATTCCAGGAA	GGCGGGTCCG	GGCAGCAGTA
	25981	CGGCGCCGTC	GACGGAGTGA	CCGGCCAGCC	ATGGGTGGGT	GGCCAGCGAG	AACCGGCCGG

	26041	TGAGCAGCAC	CTCGTCGGAG	TCGGGGAAAA	CCACCCACCC	GGCGAGCAGC	CCCTCCTCCA
	26101	TOGGGGTCGAG	TOOGAGGCCG			GGTCTCGATC	
	261.61	CATGGTGGAA	GGCGTATGTG		GTGCCGTCGC	CGTCGCGGGG	
	26221	COCAGTOGAC	GGGCACGCCG	GTTGTGTGCG	COTOGGCCAG	CGCGGTGAGC	
5	16281	070000000000	GOGGOGGAGO	3773737603 3733663A063	T0303306T0	GATOGOGGGG	
5	36341	3376060667	GACCTCGACG	AACACGGTGT	- CACCCGGGCTC	GEGGGCAGCG	· ·
	26401	TGGCGAAGCC	TACGGGGTGG	CGCATGTTGC	GGAACCAGTA	CTCGTCGTCG	
	26461	CGATCCAGCG	TTCGTCGGCG	GTGGAGAACC	ACGGGGATOIC		GAGGTGGTGT
	26501	COGCGACGAT	CCGCTGGAGT	TOGTOGTACA	-900000100A0	GAACGGGGTG	TGGGTCGGGC
10	26581	ASTOGACGAS	GATGCGGCGC	ACCCAGACGC	- adadaaaaawa - adadaaaaawa		ATCAGCGTTT
• •	26641	CGACGGCGTC	2G3GCGCCCG	GOGAOGGTOG		GCCGTTGCGG	CCCGCGACCC
	26701	AGACGCCGTC	GATCCGGGCG	GCATCCGCCT	221 10:0000		GCGACCGAGC
	26761	20ATCG0S0C	GCGTCCGGCG	AOTTCGCGCA	oumos, oma men inneern	AACGCTGCGC	
	26821	GGCGGGGACC	GTCCTCCAGG	GTGAGCGCTC	23303A3A0A		ATCTCGUCCT
15	26881	GGGAGTGTCC	GATGACGGCG	TOCGGGGGGTA	-03000300000 -03000330230		GOGGCCAGCG
	26941	ACACCATGAC	GGCCCAGCAG	ACGGGGTGCA	COACCTCGAC		ACCTCCGGGT
	27001	CGTCGAGCAT	GGCGATGGGG	TOCCAGOCOG	TOTOCHORA		GCGCATTGGC
	27061	GUATOOTOGO	GGCGAACACC	GGGGAGGCCG	0010010000	GAGGCCCATG	CCGCGCCACT
	27121	GCGGTCCTTG	TOOGGGGAAG	ACGAAGACGG	TG000000000	GGTGAGCGCC	GTGCCGGTGA
20		CGACGTCGTC	GTOGAGCAGC	ACGGCGCGGT	3033354037	CGTACGCCTG	SCGAGCAGGC
~0	27241	CCGCGGCGAT	GGCGCGCGGG	TCGTGGCCGG	38333333333333333333333333333333333333	GAGGTGCTCG	CGGAGTCGGC
	27301	GGACGTGGCG	GTOGAGGGCC	GTGGCGGTCC	G03000A0A0	GGGCAGTGGT	GTGAGCGGCG
	27361	TGGCGATCAG	CGGCTCACCG	GGCTTCGAGG	COGREGACTE	CTCGGCCGGC	GGCTCCCCGG
	27421	00000M10M0		ACGTGGGCGT	Tegraces	GACGCCGAAG	GAGGACACAC
25	07481	CGGCGCGCCG	CGGGCGGTCG	GTCTCGGGGCC	AGGGGGGGGG		AGTTCGACGG
~ 3	27541	CGCCGGCCGT	CCAGTCGACG	TGCGAGGACG	GOGTGTCCAC	GTGCAGGGTG	CGCGGCAGGG
	27601	TGCCGTGCCG		ACCATCTTGA	TGACACCGGC		GCGGCCTGAG
	27661	TGTGGCCGAT	GTTGGACTTC	AGOGAGOCCA	GCAGCACCG	GGTGTCGCGC	CCCTGCCCGT
	27721	AGGTGGCCAG	CACCGCCTGT	GCCTCGATGG	GATOSCOCAG		GTGCCGTGCG
30	27781	CCTCCACGGC	GTCCACGTCC	GCCGGGGTGA	GCCCGGCGTT	GGCCAGGGCC	TGCCGGATCA
50	27841	COCGCTCCTG	CGAGGGCCCG	TTCGGCGCCG	ACAACCCCTT	GGAAGCACCG	TCCTGGTTGA
	27901	CCGCCGAACC	CCGGACAACC	GCCAGCACAC	GGTGGCCGTT	GCGCTCGGCA	TOGGAGAGCO
	27961	TCTCGACGAT			CGAAACCGGT	GCCGTCAGCC	GCATCCUCGA
	28021	ACGCCTTGCA		GGCGCGAGAC	CCCGCTGCTG		ACGAAGCCGG
35	28081	ACGCCGAGGC		ACGCCGCCGA		CGAGCATTCG	CCGGAGCGCA
	28141	GTGACTGCCC	GGCCTGGTGC		GOGACGACGA		TCGACCGTGA
	28201				ACAGOOGACC	GGACAGCACA	
	28261	TGCCGGTCGC	GCCGAAACCG	CCCAGGTCGG		GTACCCGTCG	GAGAAGGCGC
	28321	CCATGAACAC		CTTCCGCGCA			GCGTGTTCCA
40	28361		CGAGGTCTCC			GTCCATCGCC	
	26441	SCGGACTGAT					
		GACGCACGGT					
		AACCACGGTC					
		AGTOCTOCGG					
45		SCTCGTCCTG					
		GCGCCGCGGT					
		CGGGCAGCCG					
	28861	AGTCGACGCC	GAGTTCCTTG	AACGTGGCGG	TOGCCTCGAC	CCGTGCGGCA	CCGTCGTGGC
	28921	CGAGTACGGC	CGCGGTGCAC	TGCCGGACGA	CGGCGAGCAC	GTCCTTTTCG	GCGTCCGCGG
50	28931	CGGAGAGCCG	CGCGATCCGG	TCGGCGAGGG	TGGTGGCGCC	GGCCGCCCGG	CGCCGCGGCT
	29041	CCCGGCGCGG	TGCGCGCAGC	AGGGGCGAGC	TGCCGAGGCC	GGCCGGGTCG	GCGGCGACCA
	29101	GCGCCGGGTC	CGAGGACCGC	AACGCCGCGT	CGAACAGCGT	CAGTCCGCCT	TCGGCGUTCA
	29161	GCGCCGTCAC	GCCGTCGCGG	CGCATGCGGG	CGCCGGTGCC	GACCGTCAGC	COGCTCTCCG
	29221	GTTCCCACAG	GCCCCAGGCC	ACGGACAACG	CGGGCAGTCC	GGCTGCCCGG	CGCTGTTCGG
55		CCAGCGCGTC					
		CACCGGCGGC					
		GCAGGTGCCA					
		GCGCGGTGAG					
	29527	GCGCCGGGTC	GATCCCCGCC	AGTACGGAGG	CGAGTTCGTC	CCGGTCGGCG	ACGTCGCAGG
60	29581					GCCGCTGCGC	
	29641	GCAGCCGGCG	CACGCCGTGG	CGTTCGACGA	GGTGGCGGCT	GATGATGCCG	GCCAGCGTCC
	29701	CGGAGCCACC	GGTGACGAGC	ACGGTGCCGT	CCGGGTCGAG	CGCCGGAGCG	TCACCCGCCG
		GGACCGCCGG					
	29821	CATCGAGCGC	GGTGGCCGCT	GCGAGCAGCG	GCTCGGCGGT	GTCCGGGGCG	GCGTCGACGA

	00004			maccoancac	000000000000000000000000000000000000000	G1.0m000000	
		GGACGATCCG	GCCGGGGTGT				
		ACCCGAGACC	GGGCCGG-1-2	TGGACGGCCA	015AUUUUU10	GSCGTACCGG	
	30001	GGAAGCGCTG		AGGACGCCGG	Caldunation	GCGGGTGTCG	
=		CACCCCCCCC	GCCGTGCGCG	GGGAGGATCA	COACGTCCGG	GACCGTCGGG	
5	3012.	33000371337	000000000000	9909GCAGCT	000000000000000000000000000000000000000	3300AGCACC	
	30.5	GGCCCGGAAC	GGGTCCCGTG	ACCGTCAG33	GGCGCCTGCG	CACGGCGCCG	
	30241	00000000000	ggreregree	GOGAGGTOTA	COTOGAACGC	GGTGACGGCG	
	30301	COGTGGCGCC		ACGCGGACGT		GTACGGAAGG	
10	30361	CCGCGGCGAG		GOGOCGAGCA	GCGCCGGGTG	CAGGCCGTAC	
10	30401	CGGCGAGCTG	TCCGTCGGCG	AGGGCCACTT	COGCCCAGAC		TOGGCCCAGA
	30481	CBGCGCGCGG			2000000000	GRETCGGGCC	
	30541	COATGTCGTC	GGGGTCCACC	9300663006	03A06ATT10	CCACGTCGAC GTGCGCGTGC	
	30601	3CACGGCCGG		SSGTCGSGGG	0000000000000	3000000000	
15	30661	0000000000		TORCURUS. NA	ACCCCCCCCC		
10	30721	TCACCGTGAC			CGTCGCCCCC	COTGAGGGGG	DESTRUCTION
	30781	TGAACGTGTC		CAGCCGGCTT		COGGATOGOO	
	30641	GGGCCGCGGC		GCGAGGGGG	GCAGGGASTS TGCCGGGCAG		
	30901	CGACCCGGCC		AGGTCGCCGG	GGGCCGCSTC	GGTGACCGCC GCCCGCGGTC	GOGGTCAGCG
20		CCGGGTGCGC		TGTCCGGCCG AACGGGTACG	TCGGCGGGTG		
20		GCCAGTAGCG		TCGACCGTGA		OGAGGCGCGT OTGCAGCCGG	GOOGGOGGGG
		GGTCGATGAC			GCAGCATOGG		
	31141	TCAGGGCGGA		TOGTOGGCGT	SCACOGOCCC	GATGCCGTCG GTCGTGCGCG	
		TCAGGCTCCG	GACGGTGTCG	ATCTCCAGGA		0T00G3CGTG	GOGACCTGTT
25	31261		CATCGGGATC			GOTOTOGGGG	
-2			CATCGGGATC				CGCAGGCGGG
			GAGCCGGGCC			CTCGTCACCG	
			CCCGTTGACC		CCACGCCGTC	CCGCAGCAGC	
			CGCGATCACG		COCCGCCGGA		TGCATCAGGC
30			GGACACCAGC				GCGACGTACG
50			CTCGCCGATC			CGGGCGTACG	
			GCCGAGTGCG		CGAACACCGC	GGGCTGGGCG	
			GAGCCCGGCG			CACCTOGOGG	
			GTAGGCGGCG				
35			CCACACGAGG			GCCGGTGACC	
55			GGCCCGGTGC				
		GCTCGTCCTC			CGCGCAGGCG	STGTACCTGT	
		CCTGCGGGGT				GGTGTCGGGT	
		GTTCGGGGGC		TGGCTTTCGA		AGCGTTGGTG	
40			CACCCCGGCG				
10	30021	TGAGGAGTTC	GACGGCGCCG	GCCGTCCAGT	CGACGTGCGA	GGACGGCGTG	TOCACGTGCA
	32341	GGGTGCGCGG	CAGGGTGCCG	TGCCGCATGG	CGAGGACCAT	CTTGATGACA	COGGOGACGC
			CTGAGTGTGG				
			CCCGTAGGTG				
45			ATGCGCCTCG				
			GATCACCCGC				
			GTTGACCGCC				
			GAGCCTCTCG				
			CGCGAACGCC				
50			GCCGGACGGC				
			GCGCAGCGAC				
			CGTGACCGCC				
	33001	GCACACTGGT	CTGGGTGCTG	GTGGCACCGA	AACCGCCGCG	GTCGGCTCCA	GTGCCG/ACC
	33061	CGTAGAAGTA	GCCGCCCATG	AACACGCCGG	TGTCGCTTCC	GCGCAGCGAC	TCCGGGAGGA
55	33121	TCCCGGCGTG	TTCCAGCGCC	TCCCACGAGG	TCTCCAGGAC	CAGACGCTGC	TGCGGGTCCA
			CTCACGCGGA				
			ACCATGACGC				
			GTCCCAACCA				
			CCACAAGTCC				
60	33421	CCACGATCGC	CAACGGCTCG	TCCTGCCGGA	CGGCCGCGGT	CGGGGTACGC	CGCCGGGTGG
			GCCGGCCAGT				
	33541	CGAAGACGAG	CGTAGCGGGC	AGCGTCAGGC	CCGTCGCGTC	GGCCAGCCGG	TTGCGCAGTT
	33601	CGACGCCGGT	CAGCGAGTCG	AAGCCCACTT	CCCTGAACGC	GCGCGCGGGT	GCGATGGCGT
	33661	GGGCGTCGCG	GTGGCCGAGC	ACCGCGGCAG	CGCTGGTACG	GACGAGGTCG	AGCATGTCGC

						GAGGGTGCGT	
	33781	GGACCCGGTC	GGACGCGGCG	ACGGCGGCGA	GGTCGAGCCG	GATCGGCACG	AGCGCGGGCC
		SSTOGSTGTG				TGCGGCCGTC	
	33901		GGCGATGCGG				
5							
J	33961	ACCORDORAGE	CAGTOCOCAG			CCCCTGGTGG	
	34021	GGGGGAGCGC	GTCGAGGAAC	GOGTTGCCGG	TOGOGTAGTT	GGCCTGACCC	
	34081	ACCTGCCGGA	TATGGACGAG	TACAGGAIGA	#090000073	GTOGAGATOG	CGCGTCAGCT
	34141	CGTGCAGGTG	CCAGGCGACG	TOOGGOOTTGA	CCCGCAGCAC	GGCGTCCCAC	TGCTCCGGCC
	34201	GCATGGTCGT		TOSTOSACSA		GTGCACGACG	
10		GCTGGGCGAC		ACTGCGGCCA		GTCGACGACG	
10		CGTACCGCAC			03000333000	GCCGTTGCGG	
			GCGGTCGTCC				
		CGACCTCGGC		ACGGTGAGCA		GAGGAGGCGG	
	34441	CGGTGCCGCC		ACGGTCCCGC		GGAGGTTCCG	GTGGCCGCGG
	34501	COMOMOGGGG	CAGACGGGCC	GCACGCGCTG	TBCCGTCGGC	GACCCGGACG	TGCGGCTCGT
15	34561	000000000000	GAGCCCGGCC	GCTATGGCGG	CGGGGGGTGAT	CICGICCGCT	TCGATCAGCG
	34621		GGGATGCTCC			BOOGCOGAGO	
	34681		GGTACGGGTG			CGCCCAGCGC	
					GGCCCAGCTC	COGGGTCCGG	
						1000010000	
2.0		AGGTGCCCGG		TOCAC JA	GGACCACGAC	CGGGGGGTGC	
20	34861		GAGGTACGTC	CAGTCGGGGA		GGGCACGGGC	ACCCAGGCGA
	34921	TOTOGRACAG	CGCCTCGGCA	TOGGGGTTTGG	CGGCCCGCAC	GGTCAGGCTG	TCGACGTCAA
	34981	GGACCGGTGA	GCCGTGCTCG	TCCGTGGCGA	CGATGCGGAC	CATGTCGGGG	CCGACGCGTT
	35041	CCASCAGCAC	GCGCAGCGCG			CCTCACGCCG	GACCAGGAGA
		ACGCCAGCCG		GGGTCCGTGA			
25		ACCOMMODE	000000000000	0000000000	COROCOTICO	CTGTTCGGCG	
23	35161						
			GCGGCCCTCC				
						CGGGTCGGCG	
	35341	GCCAGTCCAC	GGGCTCCGCC	GGACCGCCAG	TGTCCACGCT	CAGCGCTCCG	GTCGCACTGA
	35401	GCGCCCAGGG	GCCCGTGCCG	GTACGGCTGT	GCAGACTCAC	CGACCGCCGT	CCGGACACCT
30						GTCGACCACC	
						TCCCGCTTCG	
	35521	007.7007.000	COLCOCCOCC	7000000730	ccaccmeate	CTCGTGGTCG	GCGAGCCAGG
						TCGCCGTCGG	
2.5			GAGCAGCGGG			TTCCGCGTCG	
35						ACGCGTCGCG	
	35821	TGACGGGCAC	GCCCCGGACC	CAGAGCGCGG	CGAGCGACCG	AGTGAAGCGG	TCCAGGCCGC
	35881	CCTCGCCTCG	CCGCAGTGTG	CCGGTGACGA	CCGTATGCGC	ATGCCCGGCG	AGCGTGTCCT
	35941	CCAGTGCGGT	GGTGAGCACG	GGATGCGCGC	TGACCTCGAC	GAACGCGCGG	TATCCGCGGT
						GCGCAGGTTG	
40						GGTGGAGAAG	
40	30001	A33333CG1C	000000000	CCASCAGO	CCTCGTCCAC	COCCOCCCC	A TICOMOTOCO A
	30121	Company	CGOMGTGATG	COGGGAGAG	CG.CGMGCMG	CGCGCCGCGG	ATCOITICGA
						GGCGCGGGGG	
						CGACGCGGGT	
	36301	CGGCGACCTC	CAGGCGCCCG	GCCCACACGG	CGGCGTCGAA	GTCGGCGGGC	GGCACCGAGA
45	36361	CCATGCCGCC	CTGCCCGGCC	AGTTCGGTGG	CGACGAGTCG	GCTGCGCACC	GCGACGACCT
						CGCGGCGACT	
						CCACAGCTCC	
						GTCGAACGCG	
50						CGGGGCGAGC	
50						CAGTTCGGCA	
						TGTGTCGGTG	
	36781	TTCCCGTCAC	GGCCCCCGGC	ACTTCGGCAC	CACGGGGGAA	CGCCTCCGCC	TCTCGGGCCG
	36841	GCACGACCGC	COGGTGGCGC	ATGGCCGTCC	GGGTGGTGGC	GAGCGAGTGG	CCGACCGCGG
						CCGCAGTCCC	
55						GGCTTCGGGT	
55							
						GGCGTTGGTG	
						CGGCCACGGC	
						CGGCTCGTCG	
	37201	TGCGCGGCAG	GACGCCGTGC	CGCATCGCCA	TGACCATCTT	GATGACGCCG	GCGACGCCGG
60	37261	CCGCGGCCTG	GGTGTGGCCG	ATGTTCGACT	TGAGCGAGCC	GATCAGCAGC	GGATGCACGC
	37321	GTTCGCGCCC	GTAGGCCACT	TGCAGGGCCT	GGGCCTCGAC	GGGGTCGCCG	AGACGGGTGC
	37301		011100000000	GCGTCGTCGT	CACCGGGGGG	CAGGCCGGCG	TOGGOGAGOG
	21201	01000000000	TOUCTUCACO	40000000000000000000000000000000000000	CCEMCCCCCC	CHARGECAGCA	TTCC20001100
	3/441	CAUGUTGGAT	GAUGUGCTGC	20000A00001		GGACAGCCCG	TICGMCGCGC
	3/501	CGTCGGAGTT	GACCGCGGAG	CCGCGCACCA	GUGUUAGUAC	GGGGTGGCCG	1000000100

	22.5.5.		_				
		CGTCGGAGAG					
		CGGTGTCCGC	SAAGGCCTTG	GCACGGCCGT	000000000000000000000000000000000000000	COCGCGCTGC	CGGGAGAACT
	37681					GACCAGGGCG	
5	3/41	CCCCGAGCG	CAGCGACCGG	GCGGCCTGGT	GOAGOGCCAT	CAGCGACGAC	GAACACGCCG
.)	37801	TGTCGACGGT	GACCGACGGG	CCCTCCAGAC	CGAAGTAGTA	CGAGAGCCGC	CCGGAGAGAA
		COCTGGTCGG					
	3/921	GGGTGAACGC	SCCOATGAAT	ACGCCGGTGT	DGGTGGGGG	PACSCITICS	GGCAGGATGC
	37981	CCGCTCGTTC	GAACGCCTCC	CACGACGCTT	CGACCAC	ACGCTGCTGC	GGGTCCATCG
1.0	38041	CCAGCGCCTC	ACGCGGGCTG	ATCCCGAAGA	Accordacto	GAAGTCGGCG	
10	38101	GGAAGCCGCC	GTGACGCACG	GAAACCTTGC	CGACCGCCTI		
		CGGCGAGGTC				CGCGTCCCCG	
		COAGCOGCOA					
	38281					CACTGTCGCC	
15	2034.	02000000000	CICACOCCGC	COTTUCTOA.	Compagning	TOUGHGUGUG	GUCGGTGTCG
Į J	- 2040± - 50355*	GGTGGTCGAA	GAUGGCCUTC		OSTOSTOST	Laru.CaGCG	AGGCTGTTGC
	30501	GCAACCGGAC	ACCULTGAGO	GASTUGATOU	- Comballia	UNACGUCGTU	GI GGGCGTGA
	30521	TCTCGGAGGC	CGGCCCC	CCCMGCMCCC	030000.000	COUMUACACO	
		GGTCACGATG					
20	38641		CTGCCGGACG			GACCATGAAC	
20		CGGGGAGGCT					
		ACATGCCCCA CGTCGAGGAA					
	350.1	CGGCGCTGGA	COCG. 16565	T. COCCOCATA	00000000000		MOSMOCACOM
25	30001	GCCAGGCGGC	CEMCCCEEEC	CCCCCCCCCC	0000.100	7.1.23010A00	20010CAGG1
ديد		CGAGGATGCC					
		TGTGGGCGAG					
		CGGGGGTGGT					
		GGCGGGCGAG					
30		GGTTGAGGGG					
		GGAGGGTGTG					
		GGAGGGGAGT					
		GGGCGGTGCG					
	39541	TGAGGGTGTG	GTCGGTGGTG	GTGAGGTGGT	STTSCAGGGC	GGTCAGGACG	CGGGTGGCGC
35		GGGTGTGGGC					
	39661	CGTGTCCCTC	GGGCAGGTCA	CCGTCGTAGA	CCGCCTCGGC	GACCGCGAGC	CACTCCAACC
	39721	GGAGCGGGTT	CGGCCCCGAC	GGGGTGTCGG	CCCGCTCCCT	CAGCACCAGC	GAGTCCACCG
		ACACGACAGG				GACGCCGGCC	
		TGAGGGGGAC					
40		ACGGCAGCTC					
		GTGCCGGATG					
	40021	CGGCATACAC	GGTGTCACCA	TCACGCCAGG	CAGCCCGCAA	CCCCTGGAAC	GCCGACCCGT
		ACTCATAACC					
		TGGCCGGCGG					
45		GGGTCAGGGT					
		CGGTCACCGG					
		CTGCGGTCAC					
		CGGTCTCGTC					
•		TGCCCCGCAC					
50		GAACAACACC					
		CCGCCCCGGT					
		GCCTGTGCTC					
		CCGTGCCCCA					
c c		GCTCCCAGCC					
55		GCAGCAGCAC					
		CCGCATCCAG					
		CGG1'CACCCA					
		TTCCCTTCAG					
<i>(</i> 0		CGTAGTCGAC					
60		CCTCCACCGC					
		TCCACACACC					
		CCCGGCCGGC					
		CCTCCAGGCT					
	41341	CCACAGCGTC	CGGCACGACC	CCATGCGCCT	GCCACAGCGC	GGCCAGGCTC	ACCGCGACCG

	41401	and corece	coordeater	indrocanna	COTTOCCOCAC	ATCCGACCGC	CACAACATCT
		CCCGCACATC				ACACTCCTCC	
	41521	OGARCACOGO		ATGAGTTCCA	CGCCCATGCC	CACCCACTGG	GCACCCTGCC
	41581	CGGGGAAGAC	GAACACCGTA	CGCGGCTGAT		ACCCATCACC	CGGGCATCAC
5	41641	COAGCAGCAC	CGCACGGTGA	nngaagacan	CACGGGGGAGG	CACCAACCCC	TGCGCGACCG
	4175	233222222	CLCCCCACCC	709000A0A**	ACCCCTCCAG	COGCTCCACC	TGCCCCCGCA
	41761	720002000	ACCACGAGGG	GACACOGGCA	ACGCCACCAA		CCCGACTCCA
	41821	CACGOGACGG		COCTCCAGGA	TCACGTGCGC	GTTCGTACCG	CTCACCCCGA
	41881	ACGACGACAC	ACCCGCATGC	SSTSCCCGAT	CCGACTCGGG	CCACGGCCTC	GCCTCGGTGA
10	41941	GCAGCTCCAC	CGCACCGGCC	CACCACTOCA	CATGOGACGA	CGGCTCGTCC	ACGTGCAGCG
	42001	TOTTOGGOGG	GATCCCATGC	CGCATCGCCA		GATGACACCG	GCGACACCCG
	42061	CAGCCGCCTG		ATGTTCGACT	TGACCGAACC	GAGGTAGAGC	GGCGTGTCGC
	42121	GGTCCTGCCC			GOGCCTCGAT		AGCCGCGTGC
	42181	CGGTGCCGTG	CGCCTCCACC	ACGTCCACAT	oggoggogga	CAGTCCGGCG	TTGACCAACG
15	30011	CONGCOGGAN	CACGCGCTGC	TBBBBBACGE	CGTTGGGGGG	GGACAGTOCG	TTGGAGGCAC
	42301	CSTCCTGGTT	CACCGCCGAS		CCGCGAGAAC	SGTGTGCCCG	TTGCGCTCGG
	42361	CGTCGGAGAG	CCGCTCCAGC	ACGAGAACGC	CGACGCCCTC	GGCGAAGCCG	GTCCCGTCCG
	121	13000	GAACGCCTTG	CACCGTCCG"	20000004050	madadaaraa	CGGGAG. IT
	42481	CCACGAGUTC	TGCGGTGTTC	SCCATGACGG	TGACACCGCC	GACCAGCGCC	AGGG/.GCACT
20	42541	occeggees	CAGTGCCTGT	GOOGCCTGGT	GCAGGGGGAC	CAGCGACGAC	GAGCACGCCG
	42601	TGTCGACCGT	GACCGCCGGG	CCCTGAAGTC	CGTACACGTA	CGAGAGGCGC	CCGGACAGGA
	42661	CGCTCGTCTG	CGTCGCCGTG	ACACCGAGCC	CGCCCAGGTC	CCGGCCGACG	CCGTAGCCCT
	42721	GGTTGAACGC	GCCCATGAAC	ACGCCGGTGT	CGCTCTCCCG	GAGCCTGTCC	GGCACGATGC
	42751	CGGCGTTCTC	GAACGCCTCC	CAGGAGGTCT	CCAGGATCAG	GCGCTGCTGG	GGGTCCATCG
25	42841	CCAGCGCCTC	GTTCGGACTG	ATGCCGAAGA	ACGCGGCGTC	GHACCCGGCG	CCGGCCAGGA
	42901	ATCCGCCGTG	GCGTGTCGTG	GAGCGGCCGG	CCGCGTCCGG	STCCGGGTCG	TACAGCGCGT
	42961	CGACGTCCCA	GCCCCGGTCG	GTGGGGAACT	CGGTGATCGC	CTCGGTACCG	GCGGCGACGA
	43021	GCCGCCACAG	GTCCTCCGGC	GAGGCGACCC	CGCCGGGCAG	TOGGCACGOO	ATGCCGACGA
	43081	TCGCGACGGG	GTCGCCGGAG	CCGAGGGTCT	GGGCGGTCGC	GGGTGCCGCT	GTCGCGGAGC
30	43141	CGGCGAGGTG	GGCGGCGAAC	GCACGCGGAG	TGGGGTGGTC	GAACGCGGTT	GACGCGGGCA
	43201	CCCGCAGACC	CGTCCGCGCG	GCGACGGTGT		GACGGTGGTG	AGCGAGTCGA
	43261			CGGTCCGGGG			AGGCCCAGGA
	43321	CGGTGGCGAC	GCTGTCGCGG				GCACGGGCCG
	43381	CGGCGAGGCG	GTTCGCCCAC		TGGCGTCGGG		CCGGTCAGTG
35	43441		CGGCGGCGTG		TCGTCGCGGC		GCGGAACCGG
	43501	TCCGGGCCAC	GATGTACGAG		CGATGGCCTT		TCGCCGGTGA
	43561	GCGCCGGCCG	TTCGATGCCG	GGCAGCGCGC		GGTGGGGAGT	CCCTCCGCGG
	43621	CCCGTGGCCG	GGTGTGGGCG	TCGGCGCCGG		GAGCAGGACG	TGCACGAGCG
.4.0	43681		CGCGGCTTCC				GTCTCGTCGC
40	43741					CAGGACCGGC	
	43801					CCGGGCGTGG	
						CGGCAGCGGG	
						CAGGCGCGCG	
4 =						GGTCTTGCCC	
45						GAGTTCACCG	
						GGCGCTGCGG	
						TTTGGCGGGA	
						CGCCATGCCG	
50						GCCCGCGTCG	
30						GGGGAACGAC	
						CCGGAACGCG GGGTCCGACT	
						GGTGCCGAGC	
						GACCTCGCCG	
55						AGCGTTCCGG	
33						CGCGCGTACC	
						GAGCGAGGCC	
						TCCGGGTTCG	
60						CGGGATCTCG	
UU						CGGGGTGCCG	
						GCCGGTGCGG CGCGGTGCGG	
						CCGGCGGCGG	
						GTAGAGCACG	
	40101	GGAICGCCIC	ひしんじじじゅんじ		COGAMACOAC	GINGAGCACG	COLUMN

	400.4			0000000000			
					CGGCGTCCCG		
		CGACGGTCTC			TOTOCOCCECC		
		CCCGGCCGGT		TGTCCGGTCT		TECGAGGTCC	
_	4.5401	ACCAGGGGTC	CACGAGCACC	TGGGCGGTCG	CCTCCGGCTG		
5	45481	GGCTCGGCCC	GCTCGCCCAC	AGCTCGCCCT	COTOGOCGGG	TGCCACGTCG	GCGCCGGACA
	45541	10036103A0	GARCCGCAGC	GACAGGCCCG	GCACGGGCAG	CCCGCACGAG	CCGGGAACCC
	45611	GCGCATCCTC	CAGGGTGTTG	GCGGTGAGCG	AGCOGGTOGT	CTCGGTGCAG	COSTACGTGT
	45681	CGAGCAGGGG	CACGCCGAAC	GTCGCCTCGA	AATOCOTGGT	JAGCGACGCC	GGCGAGGTGG
	45721	ATCCGGCGAC	CAGCGCCACG	CGCAGCGCGC		STOGCOGGAO	
10	45781			GGCACGCCGA	CGAGCACGGT		
	45841		GTCACGCGCG			GGCGGACGCG	
					GGCTGTGGAA		
					GCACGTCGCA		
					GACGGCCGGT		
15	46081						
					0000003000		
	46201				GGGTTT0GT0		
		CGGAGTCCCT		-130103M3G.	GGTCG20G30	33.3MCCM3C	WOOGLCGCGC
					AGACCTOGAT		
20					CGCCGGACGC		
20							
					TCACGGCGCG		
					GGAGCAATTC		
					GACCAACCGC		
25					TTTCCGGACC		
23					CTAGGGGGTT		
					CCAGGGAAGG		
					GACCGGATCA		
					CGGCCACACC		
20					CCGTGACGAC		
30					GGACAACGGC		
					CGAGAGCGGC		
					CCGGTCGCTG		
					GACGAACGCG		
35					CGACGAGGGC		
33					GGGCGGCATC		
					cocceccerc		
					CGCACGGGCC		
					CGGCATCCTC		
10					CCTCGACAAC		
40					GGGGCACGTC		
					CGATACGGTC		
					GCACACCGAC		
					ATACCCGGTA		
					GTCACCGCAC		
45					CTCCAGCAAT		
					GGTCTATTGG		
					GCGTTGCGGA		
					CCGGAACACC		
- 0					CGCGACGAGA		
50					ATCCGCCTCG		
					GGCTGGTCGT		
					ACTGCCCGCC		
					GAGCGGCGCG		
					CTCCGGGGGG		
55					GACGCGGACG		
					ACGCTCGCGC		
					GTCCTCGCCC		
	48661	GTGCTGGTCG	GCACGCCCGT	GGCGAACCGT	ACGCGGGCGG	CGTACGAGGG	CCTGATCGGC
	48721	ATETTCGTCA	ACACGCTCGC	GCTGCGCGGC	GACCTCTCGG	GCGATCCGTC	GTTCCGGGAA
60	48781	CTCCTCGACC	GCTGCCGGGC	CACGACCACG	GACGCGTTCG	CCCACGCCGA	CCTGCCGTTC
	48841	GAGAACGTCA	TCGAACTCGT	CGCACCGGAA	CGCGACCTGT	CGGTCAACCC	GGTCGTCCAG
	48901	GTGCTGTTGC	AGGTGCTGCG	GCGCGACGCG	GCGACGGCCG	CGCTGCCCGG	CATCGCGGCC
	48961	GAACCGTTCC	GCACCGGACG	CTGGTTCACC	CGCTTCGACC	TCGAATTCCA	TGTGTACGAG
	49021	GAGCCGGGTG	GCGCGCTGAC	CGGCGAACTG	CTCTACAGCC	GTGCGCTGTT	CGACGAGCCA

	10001	000000000000	COMMOGMOCT	Coremonares	COCCUACTOC	- CCCCCCCCA	acaean acae
			GGTTGCTGGA				
		GACGTACGGC			GACGCGACGG		CGTGGTGCCC
	49201	TOGAROGACA	CGGCGCGGGA	CCTGCCCGTC	GACACGCTGC	CGGGCCTGCT	GGCCCGGTAC
_	49261	GCCGCACGCA	CCCCCGGCGC			ACATCTCCCT	CACCTACGCG
5	49321	CAGCTGGACC	GGCGGGGGAA		CACCIGCTCC	GCGCGCGCGG	CACCGCCACC
	49351	GGCGACCTGG		CBCCGATCGC	GGCGCCGACC	TGATCGTCGG	CATCGTGGGG
	49441	ATCCTCAAGG	cagacacacac	TTATGTGCCG	CTGGACCCC	AACATCCTCC	GGAGCGCACG
	49501	GCGTTCGTGC	TGGCCGACGC	GCAGCTGACC	ACGGTGGTGG	CGCACGAGGT	CTACCGTTCC
	49561	coeffeedes	ATGTGCCGCA	COTOSTGGCG	TTGGACGACC	CGGAGCTGGA	CCGGCAGCCG
10	49631	GACGACACGG		CGAGCTGGAC	CGGGACAGCC	TOGCCTACGC	GATCTACACG
	49581	TCCGGGTCGA	CCGGCAGGCC	GAAGGCCGTG	CTCATGCCGG	GTGTCAGCGC	CGTCAACCTG
	49741		AGGAGCGCAC		GAGCCGGCCA		CCAGTTCGTG
		ACGCCCACGT	TCGACTACTC			CGCTGCTGGG	CGGCACGCTC
	49361			GCGGTTCGAC	CCGCCGGGAC	TOGCOCGGTG	GATGGACGAA
1.5	49921		CCCGGATCTA		GCCGTACTGC		CGAGCACGTC
13		GATCCGCACA		CGCCGCCCTG	CGGCACCTST	GCCAGGGCGG	CGAGGCGCTG
		ATCCTCGACG	CGCGGTTGCG	CGAGCIGTGC		CCCACCTGCG	CGTGCACAAT
	50101				ACCGGGTACA		CCTCCCCGAC
		GOGTGGCCCG		GATOGGCCCG		ACACCCGCAT	CCATCTGCTC
20						AGCTCTGCGT	
20		GACGAGGCGA		TOOGGACGGT			CGCCGGCGTC
	50281	GGCCTCGCCC		GGCCCGTCCC	ACCGGCGACC	CCGAGCGCTG	GGTGCCGGGA
	50341	GATGCGGTCG				TGGCCCGCCG	CGCGCCCGAC
			AATTCCTCGG		GACCAGGTCA		CATCCGCGTC
2.5			AGATCGAGAG				GGCGGCGGTG
25	50521		AGGACCGGCG		TTCCTGGCCG		ACCGGTGGCC
	50581	GGCCGGCACG		CGCCGCGTCG	CTGCGCGCGG		COGGCTGCCC
						TGCCGAGGAC	CACGAGCGGC
		AAGGTGGACC			GAGCCGGGCC		CGGGGCGGTT
• •			CCGATGCCGA		TGCCGGATCT	TCCAGGAGGT	GCTCGACGTC
30			GTGCCGACGA			GGCACTCCCT	GCTCGCCACC
	50881		CCCGCATCCG			TCCCGCTGCG	TACGCTCTTC
	50941		CGCCCGCCGC				
	51001		CGCCCTCCGC				
	51061		CGCACGGCTC				
35	51121		GCGGGCCACT				
	51181	GCGCGCCACG	AGCCGCTGCG				
	51241	GCTCCGGTGC	GCGCCGAGGT	GGTTCCGGTG	CCGGTCGGCG	ACGTCGACGC	CGCGGTCCGG
	51301	GTCGCCCACC	GGGAGCTGAC	CCGGCCGTTC	GACCTCGTGA	ACGGGTCGTT	GCTGCGTGCC
	51361	GTGCTGCTGC	CGCTGGGCGC	CGAGGATCAC	GTGCTGCTGC	TGATGCTGCA	CCACCTCGCC
40	51421	GGTGACGGAT	GGTCCTTCGA	CCTCCTGGTC	CGGGAGTTGT	CGGGGACGCA	ACCGGACCTT
	51481	CCGGTGTCCT	ACACGGACGT	GGCCCGGTGG	GAACGGAGTC	CGGCCGTGAT	CGCGGCCAGG
	51541		GGGCCTACTG				
			CCGGCGGGGC				
			CGGCACGCCG				
45			TCGCCCTGGT				
			CGGACCGGGG				
			TGCGCCTCGA				
			CGATGGTGGG				
			CCGCGCTGCC				
50			GGCTGCCCGG				
2 0	52081	GACGAGATGA	CCGGCGAACT	GTCGATCAAC	CTCTTCGACG	ACGGTCGCAC	CGTCTCCGGC
			ACGATGCCGC				
			CGCTGCGTGC				
			AGCCATGCCC				
55			GAAGACCCGT				
33			CCGGCTGCCC				
	50143	A CERCOCOECTO	CGACGCCATC	T CCCCCTTCC	COTCCCTCCC	CCCCCCCCC	ACCACCGCCG
			CGGCTTCCTC GGCGCTGGCG				
60	52561	GCCCGCGCGA	GGGGGGGGG	AIGGACCCGC	AGCAGCGCCT	CCCCACTCAC	ACCICGIOGG
60			GCACGCGGGC				
			GTTCTTCCAG				
	52/41	CGAGCATTCA	CACGAGCGTG	CTCTCCGGCC	GCCTCGCGTA	COCCOCCOCC	CIGGWGGGIC
	52801	CGGCGGTCAC	GGTCGACACG	GCGTGTTCGT	CGTCGCTGGT	GGCGCTGCAC	CHCATCCCCT
	52861	AGTCGCTGCG	CTCCGGCGAA	TGCTCGCTCG	CCCTGGTCGG	CGGCGTCACG	GIGAIGGCCT

	55001			m=000000100		000000000000000000000000000000000000000	0000000000
		caccagoasa		TTOTCOGAGO		SGCCCCCGAC	
		AGGCCTTCGC		GACGGCACCG		GGGGTCCGGC	
	53041	TOGAGAAGCT		GAGGGGAACG	GCCACCGCGT		GTCCGGGGTT
<i>-</i>	53101	CCCCCTCAA	CCAGGACGGT	GCCTCCAACG	GGCTGTCCGC	GCCGAACGGG	
5		AGOGGGTGAT	COGGCAGGCC	CTGGGGAACG	COGGACTORO	CCCGGCGGAC	
	53211	TOGREGOOOR	039CA00331	ACCAGGCTGG	JOJACCOCA,1	CGAGGCACAG	
	53281		GCAGGGGGGG	GACACCCCTG	TGCTGCTGGG	STOGSTGAAG	TCCAACATCG
	53341	ACCRURUMORCA		GGCGTCGCCG	GTGTCATCAA		GCCATGCGGC
	53401	ACGGCACCCT	SCOCCGCACC	CTGCACGTGG	ACACGCDSTC		GACTGGACGG
10	53461	coccat	CGMACTCCTC	ACCGACGCCC	GGCCCTGGCC	CGAAACCGAC	
	53521	GCGCCGGTGT	CTCCTCCTTC	GGCGTCAGCG	GCACCAMOGC	CCACATCATC	CTCGAAAGCC
	53581		GGCCCCGAA	cccgccccss	CACCCGACAC	CGGACCGCTG	CCGCTGCTGC
	53641	rereggeees	CACCCCGCAG	GCACTCGACG	CACAGGTACA		GCGTTCCTCG
	53701	ACGACAACCC	CGGCGCGGAC	CGGGTCGCCG		ACTOGOCOGG	CGCACCCAGT
15	53761	TOGRECACOG	CGCCGTGCTG	CTCGGCGACA	CGCTCATCAC	CGTGAGCCCG	AACGCCGGCC
	53521	GCGGACCGGT	GGTCTTCGTC	TACTOGGGGG	AAAGCACGCT	GCACCCCCAC	ACCGGGCGGC
	53861	AACTOGOGTO	CACCTACCCC	GTGTTCGCCS	AAGCGT0300	CGAGGCCCTC	GACCACCTCG
	53941	ACCCCACCA	33300003337	ACGCACTTCC	COCHCCAGA:	CGCGCTCACC	GCGCTCCTGC
	E4001	GOTCCT GGGG	CATCACCCCG	CACGCGGTCA	TOGGCCACTO	COTCGGTGAG	ATCACCGCCG
20	54061	CGCACGCCGC	CGGTGTCCTG	TOCOTGAGGG	ACGCGGGGGGG	GCTCCTCACC	ACCCGCACUC
	54121	GCCTGATGGA	CCAACTGCCG	TOGGGGGGGG	CGATGGTCAC	CGTCCTGACC	AGCGAGGAAA
	54181	AGGCACGCCA	GGTGCTGCGG	ccccccccccc	AGATOGCCGC	COTCAACGGC	CCCCACTCCC
	54241	TEGTGCTGTC	CGGGGACGAG	GAAGCCGTAC	TOGAAGCCGC	COGGCAGOTO	GGCATCCACC
	54301	Acceptage	GACCCGCCAC	GCCGGCCACT	CCGAGCGCAT	GCAGCCACTC	GTCGCCCCCC
25	54361	TOOTOGROGT	CGCCCGGACC	CTGACCTACC	ACCAGOGGGA		CCCGGCGACC
	54421	CCACCACCGC	CGAATACTGG	GCGCACCAGG			CAGGCGCACA
	54481	CCGAGCAGTA			AGATOGGOCO	CAACCAGGAC	CTCTCGCCGC
	54541	TOGTOGACGS		CAGACCGGTA			CTGCACACCG
	54601	CGCTCGCGCA		CGCGGCGTCG	CGATCGACTG	GACGCTCGTC	CTCGGCGGGG
30		ACCGCGCGCC		CCCACGTATC	CGTTCCAGCA		TGGCTGCGGC
50	54721	CCACCTCCCG				GGTGGCGCAC	
		GCGCCGCGGT			GAGTCGTCCT	GACCGGCCGC	
	54841	CCTCCCATCC		GAGCACGCGG	TCGACGGCAC	CGTGCTCCTG	- · · · · · · · · · · · · · · · · · · ·
	54901					CTGCGACCTG	
35	54961	TOGTOATOGA			CGACCGGCGG	TGTGGCGGTC	
22	55021	TOGCCGAACC	CGACGACACG		CGGTCACCGT	CCACGCGCGG	
	55081	CGGGCCTGTG	GACCCGACAC	GCCGGCGGAT		GGCACCGGCA	
	55141		GGCACCCTGG	CCGCCCGCGG		GGTCGACGTC	
	55201	CCACGGACCC			ACGGACCGGG	OTTCCGGGGG	
40	55261	ACGACCGGTT				CCCCGACGAG	
40		ACGCCGCCCC					
		TGGCCGCGCT					
		GCATCCACGC					
45		GCACCGTCCG					
4)	22261	CGCGCCCGTA	CGCGGAAGGC	TUUGGTGAUG	GCCTGCTGCG	CCCGGTCTGG	ACCUAGCIGC
	22621	CGATGCCCGT	UCCGTCCGGG	GACGATCCGC	maragadaa		CACCCCCCACC
		ACGGCGACGT					
	55/41	GCCACCTGTC	CGCCGCCGAG	GACACCACCI	regregation	CALCOGGACC	GGCCCGGCCG
50	55801	CTGCCGCCGC	CGCGGGTCTG	GTCCGCTCGG	CGCAGGCGGA	SARCUCUGGU	CGCGTCGTGC
50	55861	TCGTCGAGGC	GTCCCCGGAC	ACCTCGGTGG	AGCTGC1CGU	CGCGTGCGCC	GCGCTGGACG
		AACCGCAGCT					
		ACCCCGCGCA					
	56041	CCGGCACGTT	GCACGACGTC	GCGCTCATAG	CCGACGACAC	GCCCCGGCGG	GCGCTCGAAG
	56101	CCGGCGAGGT	CCGCATCGAC	GTCCGCGCGG	CCGGACTGAA	CTTCCGCGAT	GTGCTGATCG
55		CGCTCGGGAC					
		AGACCGGGCC					
		GCGGCATCGG					
		GGAGCTTCAC					
		TCGACCTCGG					
60	56461	TCGGCATGGC	CGCCGCACAG	ATCGCCCGCC	ACCTGGGCGC	CGAGCTCTAC	GCCACCGCCA
	56521	GTACCGGCAA	GCAGCACGTC	CTGCGCGCCG	CCGGGCTGCC	CGACACGCAC	ATCGCCGACT
	56581	CTCGGACGAC	CGCGTTCCGG	ACCGCTTTCC	CGCGCATGGA	CGTCGTCCTG	AACGCGCTGA
	56641	CCGGCGAGTT	CATCGACGCG	TOGOTOGACO	TGCTGGACGC	CGACGGCCGG	TTCGTCGAGA
	56701	TGGGCCGCAC	CGAGCTGCGC	GACCCGGCCG	CGATCGTCCC	CGCCTACCTG	CCGTTCGACC

	56761	TGCTGGACGC	GGGCGCCGAC		AGATCCTGGG	CGAACTGCTC	CGGCTGTTCG
	56821	ACGCGGGGGGC	GCTGGAGCCG	CTGCCGGTCC	GTGCCTGGGA	CGTCCGGCAG	GCACGCGACG
	56881	CGCTCGGCTG	GATGAGCCGC	GCCCGCCACA	TOGGCAAGAA	CGTCCTGACG	CTGCCCCGGC
	56941	CGCTCGACCC	GGAGGGCGCC	STCGTCCTCA	CCGGCGGGTC	CGGCACGCTC	GCCGGCATCC
5	57001	TOGOCCGCCA	CCTGCGCGAA	COGCATGTCT	ACCTGCTGTC	COGGACGECA	CCGCCCGAGG
	57061	GGACGCCCGG	CGTCCACCTG	CCCTGCGACG	TCGGTGACUG	GGACCAGCTG	GCGGCGGCCC
	57121	TGGAGCGGGT	GGACCGGCCG	ATCACCGCCG	TGGTGCACCT	CGCCGGTGCG	CTGGACGACG
	57181	GCACCGTCGC	GTGGGTGAGG	COCGAGCGTT	TOGACACGGT	GCTGCGCCCG	AAGGCCGACG
	57241	GCCCCTGGTA	COTOCACGAS	CTGACGAAGG	AGCAGGACCT	CGCCGCGTTC	GTGCTCTACT
10	57301	CSTCGGCCGC	CGGCGTGCTC	GGCAACGCCG	GCCAGGGGCAA	CTACGTCGCC	GCGAACGCGT
	57361	TOUTEGACGE	GCTCGCCGAG	CTGCGCCACG	GTTCCGGGGCT	GCCGGCCCTC	TCCATCGCCT
	57421	GGGGGCTCTG	GGAGGACGTG	AGCGGGGTCA	CCGCGGCGCT	CGGCGAAGCC	GACCGGGACC
	57481	GGATGCGGCG	CAGCGGTTTC	CGGGCCATCA	CCGCGCAACA	GGGCATGCAC	CTGTACGAGG
	57541	cagccagaca	CACCGGAAGT	CCCGTGGTGG	TOGOGGGGGG	GCTCGACGAC	GCGCCGGACG
15	57601	TGCCGCTGCT	GCGCGGCCTG	CGGCGGACGA	CCGTCCGGCG	GGCCGCCGTC	CGGGAGTGTT
	57661	CGTCCGCCGA	CCGGCTCGCC	GOGOTGACCG	GOGACGAGGT	CGCCGAAGCG	CTGCTGACGC
	57721	TOGTCCGGGA	GAGCACCGCC	gccgrgcTcG	GCCACGTGGG	TGGCGAGGAC	ATCCCCGCGA
	57781	OGGCGGCGIT	CAAGGA DOTO	GGCATCGACT	CGCTCACCGC	GGTCCAGCTG	CGCAACGCCC
	5784	TCACCGAGGC	GACCGGTGTG	CGGCTGAACG	CCACGGCGGT	CTTCGACTTC	CCGACCCCGC
20	57901	ACGTGCTCGC	CGGGAAGCTC	GGCGACGAAC	TGACCGGCAC		GTCGTGCCCC
-	57951	GGACCGCGGC	CACGGCCGGT	GOGGACGACG	AGCCGCTGGC	GATCGTGGGA	ATGGCCTGCC
	58021	GGCTGCCCGG	CGGGGTCGCG	TCACCCGAGG	AGCTGTGGCA	CCTCGTGGCA	TCCGGCACCG
	58081	ACGCCATCAC	GGAGTTCCCG	ACGGACCGCG	GCTGGGACGT	CGACGCGATC	TACGACCCGG
	58141	ACCCCGACGC	GATCGGCAAG	ACCTTCGTCC	GGCACGGTGG		GGCGCGACAG
25	58201	GCTTCGACGC	GGCGTTCTTC	GGCATCAGCC	CGCGCGAGGC	CCTCGCGATG	GACCCGCAGC
	58261	AGCGGGTGCT	COTGGAGACG	TCGTGGGAGG	CGTTCGAAAG		ACCCCGGACT
	58321	CGACCCGCGG	CAGCGACACC	GGCGTGTTCG	TCGGCGCCTT	CTCCTACGGT	TACGGCACCG
	59381	GTGCGGACAC	CGACGGCTTC	GGCGCGACCG	GCTCGCAGAC		TCCGGCCGGC
	58441	TGTCGTACTT	CTACGGTCTG	GAGGGTCCGG	CGGTCACGGT	CGACACGGCG	TGTTCGTCGT
30	58501	CGCTGGTGGC	GCTGCACCAG	GCCGGGCAGT	CGCTGCGCTC	CGGCGAATGC	TOGOTOGOCO
	58561	TGGTCGGCGG		ATGGCGTCTC	CCGGCGGCTT	CGTGGAGTTC	TCCCGGCAGC
	58621	GCGGCCTCGC		CGGGCGAAGG	CGTTCGGCGC	GGGTGCGGAC	GGCACGAGCT
	58681	TCGCCGAGGG	TGCCGGTGTG	CTGATCGTCG		CGACGCCGAA	
	58741	ACACCGTCCT			CGGTCAACCA	GGATGGTGCC	TCCAACGGGC
35	58801	TGTCGGCGCC	GAACGGGCCG	TOGOAGGAGO	GGGTGATCCG	GCAGGCCCTG	GCCAACGCCG
	58861	GGCTCACCCC	GGCGGACGTG		AGGCCCACGG	CACCGGCACC	AGGCTGGGCG
	58921	ACCCCATCGA	GGCACAGGCG	GTACTGGCCA	CCTACGGACA	GGAGCGCGCC	ACCCCCCTGC
	58981	TGCTGGGCTC			ACGCCCAGGC	CGCGTCCGGC	
	59041	TCATCAAGAT	GGTGCAGGCC	CTCCGGCACG	GGGAGCTGCC	GCCGACGCTG	CACGCCGACG
40		AGCCGTCGCC	GCACGTCGAC			ACTGCTGACG	TOGGCOOGGC
• • • •		CGTGGCCCGA					
	59221	CCAACGCCCA	COTCATCCTS	GAGGCCGGAC	CGGTAACGGA	GACGCCCGCG	GCATCGCCTT
	59281	CCGGTGACCT	TCCCCTGCTG	GTGTCGGCAC	GCTCACCGGA	AGCGCTCGAC	GAGCAGATCC
	59341	GCCGACTGCG	CGCCTACCTG	GACACCACCC	CGGACGTCGA	CCGGGTGGCC	GTGGCACAGA
45	59401	CGCTGGCCCG	GCGCACACAC	770000000	GCGCCGTGCT	GCTCGGTGAC	ACCGTCATCA
		CCACACCCCC					
	59521	AGCATCCCGC	GATGGGCGAG	CAGCTCGCCG	CCGCCCATCC	CGTGTTCGCC	GACGCCTGGC
	59581	ATGAAGCGCT	CCGCCGCCTT	GACAACCCCG	ACCCCCACGA	CCCCACGCAC	AGCCAGCATG
	59641	TGCTCTTCGC	CCACCAGGCG	GOGTTCACCG	CCCTCCTGCG	GTCCTGGGGC	ATCACCCCGC
50	59701	ACGCGGTCAT	CGGCCACTCG	CTGGGCGAGA	TCACCGCGGC	GCACGCCGCC	GGCATCCTGT
20	59761	CGCTGGACGA	CGCGTGCACC	CTGATCACCA	CGCGCGCCCG	CCTCATGCAC	ACGCTCCCGC
	59821	CACCCGGTGC	CATGGTCACC	GTACTGACCA	GCGAAGAGAA	GGCACGCCAG	GCGTTGCGGC
	59881	CGGGCGTGGA	CATCCCCCCC	GTCAACGGGC	CCCACTCCAT	CGTGCTGTCC	GGGGACGAGG
	59941	ACGCCGTGCT	CACCGTCGCC	GGGCAGCTCG	GCATCCACCA	CCGCCTGCCC	GCCCCGCACG
55	60001	CCGGGCACTC	CGCGCACATG	GAGCCCGTGG	CCGCCGAGCT	GCTCGCCACC	ACCCGCGGGC
23	60061	TCCGCTACCA	CCCTCCCCAC	LCCTCCATTC	CGAACGACCC	CACCACCGCT	GAGTACTGGG
	60101	CCGAGCAGGT	0000236000	ATACTACTIC	DOCUMENTA	GCAGCAGTAC	CCGGACGCCG
	50101	TGTTCGTGGA	COUCANGCCC	STGCTG.TGC	TOTOCOCCOCT	CGTCGACGGG	ATCCCGCTGC
	600101	AGAACGGCAC	GWYGGGGGGG	GUCCAGGACC	TGCACACCGC	GCTCGCGCAC	CTCTACGCGC
60	60241	GCGGTGCCAC		- CTGCWCGCGC	TOGREGOTES	GTCACGGCAC	GACGCGGATG
00	60301	TGCCCGCGTA	- CC1CGWC1GG	- COCCOCATOC	DOIDDDCCOD EROTERNATOR	GTCGGCACGC	CCGGCCGCAT
	20401 10701	CCGACGCGGG	COCOLICCAM	COGCOGCACI	GMATCGCCCT	CGCCGGGTCG	CCGGGCCGGG
	60401	TGTTCACGGG	##CCC###CCC	-2.000C1CCG	ACCGCGCGGGT	GTTCGTCGCC	GAGCTGGCGC
	60541 60541	TGTTCACGGG	7100010000	WCCOGIGCGC	CGGTCGAGCC	GCTCGACATC	GCCTCCGTGC
	00541	1660060000	GUACUCUGUTC	GMC1GCGCCA	COCICOMOCO	GCICGRONIC	3001000100

			GGGCCATGGC		TACAGACCTG		
		A003003303	CCGGTTCACC	3730404000		CGCCCCGTGG	ACGCTGCACG
	€0721	CCGAGGGGGT	GCTGCGCCCC	CATGGCACGG	CCCTGCCCGA	TGCGGCCGAC	GCCGAG1GGC
	60781	CCCCACCGGG	CGCGGTGCCC	GCGGACGGGC	TGCCGGGTGT	GTGGCGCCGG	GGGGACCAGG
5	60841	TOTTOGCOGA	GGCCGAGGTG	GACGGACCOG	ACCOTTOGT	GGTGCACCCC	GACCTGCTCG
	60901	100000000	CTCCGCGGTC	GGCGACGGAA	GOOSOCAGOO	GBCCGGATGG	CGCGACCTGA
	60961	0000031011			GOGCCTGCCT	0A00033030	
		COSTOCACS	GTCGGACGCC	ACCGTACTGC			ACCGACGGAG
	61021	COATGGGATT	CGCCGCCTTC	GACGGCGCCG	GCCTGCCGGT	ACTCACCGCG	GAGGCGGTGA
	61081	CSCTGCGGGA	GGTGGCGTCA	COGTCCGGCT	CCGAGGAGTC	GGACGGCCTG	CACCGGTTGG
10	61141	ASTGGCTCGC	GGTCGCCGAG	GCGGTCTACG	ACGGTGACCT	GICCGAGGGA	CATGTCCTGA
	61201	TCACCGCCGC	CCACCCCGAC	GACCCCGAGG	ACATACCOAC	COGOGGGGGAC	ACCCGCGCCA
	61261	CCCCCCCTCCT	GACCGCCCTG	CAACACCACC	TCACCACCAC		CTCATCGTCC
	61321		CGACCCCGCC		TCACCGGCCT	CACCCGCACC	GCCCAGAACG
		AACACCCCCA		CTCATCGAAA	COGACCACCC		CTCCCCCTGG
1.5							
13	61441		CACCCTCGAC	CACCCCCACC	TCCGCCTCAC		CTCCACCACC
	61501		CCCCCTCCAC		CACCOACUAC		AACCCCGAAC
		ACGCCATCAT	CATCACCGGC	GGCTIONIU	7030033		CGCCACCTGA
	61621	ADDAGGGGA	CAUCTACCTC	CTCTCCCGGGM	000000000	TGACGCCACC	00006601,000
	61681	ACCTCCCCTG	CGACGTCGGC	GACCCCCACC	AMOTOGOGAO	CACCCTCACC	CACATOCCCC
20	51741	AACCCCTCAC	CGCCATCTTC	CACACCGCCC	CCACCCTCGA	CGACGGCATC	CTCCACGCCC
	61301		CCGCCTCACC		ACCCCAAAGO	CAACGCCGCC	TGGCACCTGC
		ACCACCTOAC			ACTTCGTCCT		GCCGCCGCCG
	61931				CCGCCGCCAA		
							GACGCCCTCG
2.5	61981				CCACCTCCAT		ATGTGGCACA
25					ACCCCGACCG		CGCCGCGGCG
	62101	GITTCCTCCC	GATCACGGAC	GACGAGGGCA	TGCGCCTCTA	CGAGGCGGCC	GTCGGCTCCG
	62161	GCGAGGACTT	CGTCATGGCC	GCCGCGATGG	ACCCGGCACA	GCCGATGACC	GGCTCCGTAC
	62221	CGCCCATCCT	GAGCGGCCTG	CGCAGGAGCG	CGCGGCGCGT	CGCCCGTGCC	GGGCAGACGT
	62281	TOGCCCAGCG	GCTCGCCGAG	CTGCCCGACG	CCGACCGCGG	CGCGGCGCTG	ACCACCCTCG
30	62341				ACGCCGACGC		
	62401				TCACCGCGAT		
	62461				CGCTGGTGTT		
	62521						
			CAAGCTCCGC			CGTGCCCACG	
2.5		CGGCACGGAC				CATGGCGTGC	
35		GCGGGGTCGC				GTCCGGCACC	
	62701	CCGAGTTCCC	CACCGACCGC	GGCTGGGACA	TCGACCGGCT	GTTCGACCCG	GACCCGGACG
	62761	CCCCCGGCAA	GACCTACGTC	CGGCACGGCG	GCTTCCTCGC	CGAGGCCGCC	GGCTTCGATG
	62821	COGCGTTCTT	CGGCATCAGC	CCGCGCGAGG	CACGGGCCAT	GGACCCGCAG	CAGCGCGTCA
	62881	TOOTOGAAAC	CTCCTGGGAG	GOGTTCGAGA		CGTGCCGGAC	
4()	62941	GRAGOGACAC	CGGCGTGTTC		TOTOCOMTOG	STACSGCSCC	GGCGTCGACC
. •	63001		CGGCGCCACC		ACAGCGTGCT	CTCCGGCCGG	
	62061				TOGACACCGC		
					CTGGAGAATG		
					ACGTCGAGTT		
45					AAGGCGCCGA		
	63301	GCGCCGGCGT	TCTTGTGCTG	GAGCGGCTCT	CCGACGCCGA	GCGCAACGGA	CACACCGTCC
	63361	TOGOGGTOGT	CCGCTCCTCC	GCCGTCAACC	AGGACGGCGC	CTCCAACGGC	ATCTCCGCAC
	63421	CCAACGCCC	CTCCCAGCAG	CGCGTCATCC	GCCAGGCCCT	CGACAAGGCC	GGGCTCGCCC
					GCACCGGAAC		
50					AGGACCGCGA		
30					CCACCGCCGG		
					CGAAGACACT		
					AACTGCTCAC		
					CGTCGCTCGG		
55	63841	ACGTGATCCT	TGAGGGTGTT	CCCGGGCCGT	CGCGTGTGGA	GCCGTCTGTT	GACGGGTTGG
	63901	TGCCGTTGCC	GGTGTCGGCT	CGGAGTGAGG	CGAGTCTGCG	GGGGCAGGTG	GAGCGGCTGG
					CCGCGGTCGC		
					TGGGTGATGC		
					CCGGGCAGGG		
60							
UU					TOGOGGOTOG		
					GGGAGATGTT		
					GGGCGGTCGC		
					TGATCGGACA		
	64381	CGGCGTGCGT	GGCCGGGGCC	CTCAGCCTTG	AGGACGCCGC	CCGCGTGGTG	GCCTTGCGCA

	51141	COOLCONOLD	0000000000	0530000000	CCCCACCCAM	GGCTTCGGTG	CCAMMCCCCC
	64511	COGGTGAGGT		01000003330	GGGGAGCGAI	GCGTAACGGC	CCCCCCCCCC
		CAGTOSTGGC	CGGCGAGCCG	magacamaa	LOGACOTOCT	GACGCGGTAT	GAGACCGAAG
	64621	GOGTGCGAGT			ADSCOTOCOA	CACGCCCCAC	GTGGAAGCCA
5	64681		ACTOSCTGAG			GAAGGCCGCG	TOGGTGGCGT
_	64741	GGTGGTCGAC		GUCTGGGTGA		GGATGAGAGT	TACTGGTACC
		GGAACCTGCG			CGGCGGTCGC		GGGTCCGTGT
	64861		CAGCGCCCAT		TGCCGGCGAT	GGAACAGGCC	CACACGGTGG
	64921	CGTCGTTGCG		GGCGGCTGGG		SACGGCGTTG	GCGCAGGCGT
10	64981	GGACCCTGGG	CGCGGCAGTG	GACTGGGACA			GGGCGGCTGC
•	65041	TOGATOTGCC				GCTGGAAGCG	GCCGGTGCCA
	65101	CCGACCTGTC	CGCGGCCGGG	CTGACAGGGG	CAGCACATCO	CATGCTGGCC	GCCATCACGG
		CACTACCCGC	CGACGACGGT		TCACCGGCCG	GATCTCGTTG	
	65221		TGATCACGCG	GTGCGGGGCA		GCCGGGCACG	
15	65281	AGCTGGTCAT	CCGGGCCGGT	GACGAGACCG		AGTGGATGAA	
		AATCCCCCCT	CGTGGTGCCG	GCGACCGCAG	CCGTGGATCT	STCGGTGACC	GTGGAAGGAG
	65401		CGGACGGCGG	CGAGTGACCG	TCCACGCCCG	CACCGAAGGC	ACCGGCAGCT
	65461	GGACCCGUCI.	CGCCAGCGGC	AUTOTGACCC	CCGACACCCC	CGACACCCCC	AACGCTTCCG
	65521	GTGTTGTCGG	TGCGGAGCCG	TTCTCGCAGT	GGCCACCTGC	CACTGCCGCG	GCCGTCGACA
20	65581	CCTCSGAGTT	CTACTTGCGC	CTGGACGCGC	TGGGCTACCG	GTTCGGACCC	ATGTTCCGCG
	65641	GAATGCGGGC		GATGGTGACA			CTCCCCGAGG
	65701	ACCGTGCCGC	CGACGCGGAC	SSTTTCGGCA	TGCACCCGGC	GCTGCTCGAC	GCGGCCTTGC
	65761	AGAGCGGCAG	CCTGCTCATG			SAGOGTGCAA	CTGCCGTTCT
	65821				GOGOGACCAT		
25	65881	CGGGCCCGGA				GAACCGTCCC	GTCGCGACGA
	65941	TOGACGOGGT		TCCCCGGAAG			CCGATGCTGC
	66001	GGGTCGGGTG				TCCGTCCGAC	
	66061	TGACGCTGCG				CCGGGACCTG	
	66121	TTCTCGACGC	GCTGCTCCGG				GGTGGCCTCG
30	66181	CCGCCAAGGC	GGCCGCAGGC				GGCCGCTTCT
	66241					GAAGCGCGAC	
	66301	CACTOGGCGA					CGGCTGATGC
	66361	GGGCCACGCC				GCAGCTGCGG	
	66421	CCGGTTCCCT	CGACGACCTT				
35	66481					CTTCCGGGAT	
			GGTCGCCGAT				GTCGTCCTGG
		AGACCGGCCC				GGTCCTGGGG	
	66661		ACCGGTCGCG			CGGCCGGATG	
. 10	66721		GCAGGCGGCG	TCCGTGATGA	CCGCGTTCGC	GACCGCGTGG	TACGGCCTGG
40		TCGACCTGGC	CGGGCTGCGC	CCTGGCGAGA	AGGICCIGAL	CCACGCGGCG	GCGACCGGTG
	66941	TCGGTGCGGC	GGCCGTCCAG	ALUGUUUUUU AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	Archaddet	GGAGGTGTAC	TCCCCCACCA
						TCTGGCCGAT	
	60961	CCGCGTTCGC	CGACGCGTTC	CCGCCGGTCG	AIGIOGISCI	CAACTCGCTC GTTCATCGAG	ACCGGIGAAI
45	6/021	TCCTCGACGC	GTCCGTCGGC	C.GUTCGCGG	TOCACOTOR	GGACGCCGGC	CCCCACCCCA
40	67081	TEGRALATUCE TO THE TERM TO THE	GUAUGUUGIU	CESCARGOCCC.	TCGACCIGAI	CGACGTGCTG	CACCCCCTCC
	67201	TGCAGCGGAT	CMICGICGAG	0190103000	COCACCCCT	CGGCTGGATG	AGCAGCGGGC
	67261	CEGICCACGC	CIGGGACGIG	COSCAGGCGC	CGCGGGCGCT	GGATCCCGAG	GGGCCGTCG
	27701	TOTAL MODEO	CAAGCIGGIG	LOGGTOG	GCATCCTCGC	CCGCCACCTG	GGCCACCCC
50	67381	DONCE TO COMPONE	CGGCTCCGGC	ACCCCACCC	CCGACACCAC	CCCCGGCACC	CACCTCCCCT
50	67441	GCGLCGTCGG	CGACCCCCAC	CLICTOSCOA	CCACCCTCGC	CCGCATCCCC	CAACCCCTCA
	67501	CCGCCGTCTT	CCACACCCCC	GGAACCCTCG	ACGACGCCCT	GCTCGACAAC	CTCACCCCCG
	67561	ACCGCGTCGA	CACCGTCCTC	PARCCCAAGG	CCGACGCCGC	CTGGCACCTG	CACCGGCTCA
	67621	CCCCCCTCAC	CARCOTOCIC	CCGTTCGTCG	TOTACTCCGC	GGTCGCCGGC	CTCATGGGCA
55	67681	40000000000000000000000000000000000000	GGGCAACTAC	GTCGCGGCGA	ACGCGTTCCT	CGACGCGCTC	GCCGAACACC
77	67741	ACCOTACACA	AGGGCTGCCC	GOGGAGTCCC	TCGCATGGGG	CATGTGGGCG	GACGTCAGCG
	67801	CGCTCACCGC	GAAACTCACC	GACGCGGACC	GCCAGCGCAT	CCGGCGCAGC	GGATTCCCGC
	67861	CGTTGAGCGC	CGCGGGGCGC	ATGCGGCTGT	TCGACGCGGC	GACGCGTACC	CCGGAACCGG
	67921	TOGTOGTOGO	GACGACCC	GACCTCACCC	AGCTCGACGG	CGCCGTCGCG	CCGTTGCTCC
60	67981	GCGGTCTGGC	CGCGCACCGG	GCCGGGCCGG	CGCGCACGGT	CGCCCGCAAC	GCCGGCGAAG
0.0	68041	AGCCCCTGGC	CGTGCGTCTT	GCCGGGCGTA	CCGCCGCCGA	GCAGCGGCGC	ATCATGCAGG
	68101	AGGTCGTGCT	CCGCCACGCG	GCCGCGGTCC	TCGCGTACGG	GCTGGGCGAC	CGCGTGGCGG
	68161	CGGACCGTCC	GTTCCGCGAG	CTCGGTTTCG	ATTCGCTGAC	CGCGGTCGAC	CTGCGCAATC
	68221	GGCTCGCGGC	CGAGACGGG	CTGCGGCTGC	CGACGACGCT	GGTGTTCAGC	CACCCGACGG
		===:500000					

	60061	000100000	22020216	COCCOOCA CO	#G:#GG:#G		-007800000
	68281	CGGAGGCGCT		CTGCTCGACC		TCCCACCGCC	
	68341	SGGAGTCCCT	3000603379	1100000000	CCGTGGCGGC	CGCGCGGGAC	
	68401	CGATCGCCAT	0010000A73	GCGTGCCGGC GGCACCGACC	TGCCCGGTGG	TGTGACGTCG	
5	68461	TOTGGGGGGT	COTTOGAGTOD	adomicuanca and	DGATCACCAC	GCCTCCTGAC	
3	68521	GGGACGTCGA	CGCGCTGTAC	GARGAGGARGE	maningangan maningangan		TACAACCTGC
	68581	GGGGGGGTA		GOGGGGAGT		STTOTTOGAC	
	68641	GOGAAGCGCT	TUGUATGGAC	CCCCAGCAAC	0001001001	CGAAACGGCG	TGGGAGGCGA
	68701 68761	TCGAGCGCGG	CCGGATCAGT	COGGCGTCGC			GTCTATGTCG
10	68321	GIGGTTCCAC	GCAGGGCTAC GAGCCTGCTG	GGGCTGGGCG TCCGGACGGC		GGAGGGCCAC	GAGGGCCCGG
10	68881	CGGTCACCGT	GGACACGGCG	TGCTCGTCGT	TGGCGTACGT	GCTGCATCTG	
	68941	GGCTGCGCCT	GGGCGAGTGC	GAACTOGOTO	TGGCCGGAGG	GGTGTCCGTA	
	69001	0360030077	OGTGGAGTTC	TCCCGCCAGC	600000000000		CGCTGCAAGT
	69061	CGTTCGGCGC	GGGCGCGGAC	GGCACGACGT	GGTGCGAGGG		CTCGTACTGG
15	69121			CGGCTCGGGC	ACACOSTSCT	CGCCGTCGTC	CGCGGCAGCG
	69191	COGTCACGTC	CGACGGCGCC	TOCAACGGCC	TCACCCCCCC	GAACGGGCTC	TOGCAGCAGC
	69241		GAAGGCGCTC	GCCGCGGCCG	GGCTGACCGG	CGCCGACGTG	GACGTCGTCG
	69301	AGGGGCACGG	CACCGGCACC	CGGCTCGGCG	Addoodatada	GGCGGACGCG	
	69361	CGTACGGGCA	GGACCGTCCG	GCACCGGTCT	GGCTGGGGCTC		AACATCGGAC
20	69421	ATGCCACGGC	CGCGGCCGGT	STCGCGGGGG	TONTONAGAT	GGTGCAGGCG	ATCGGCGCGG
	69481	GCACGATGCC	GCGGACGCTG	CATGTGGAGG	AGCCCTCGCC	CGCCGTCGAC	TGGAGCACCG
	69541	GACAGGIGIC	CCTGCTCGGC	TCCAACCGGC	CCTGGCCGGA	CGACGAGCGT	COGCGCCGGG
	69601	CGGCCGTCTC	CGCGTTCGGG	CTCAGCGGGA	CGAACGCCCA	CGTCATCCTG	GAACAGCACC
	69661	GTCCGGCGCC	COTGGCGTCC	CAGCCGCCCC	999099999	TGAGGAGTCC	CAGCCGCTGC
25	69721	CGTGGGTGCT	CTCCGCGCGG	ACTCCGGCCG	CGCTGCGGGGC	CCAGGCGGCC	CGGCTGCGCG
	69791	ACCACCTCGC	GGCGGCACCG	GACGCGGATC	CGTTGGACAT	CGGGTACGCG	CTGGCCACCA
	69841	GCCGCGCCCA	GTTCGCCCAC		TOGTUGOCAC	CACCCCGGAC	GGATTCCGTG
	69901	CCGCGCTCGA	-		AGGCGCCCGG	AGTCGTCACC	GGGACCGCTC
~ .	69961	AGGAGCGGCG			GCCAGGGCGC		GGAATGGGGC
30	70021		CCGCCGGTTC		cogococora	GGACGAGGTC	TCCGACGCGT
	70081	TOGGCAAGCA			ACGTCTACCA	CGGCGAACAC	GGCGCTCTCG
	70141		CCTGTACGCC		TGTTCACGCT	CGAAGTGGCG	
	70201		CTGGGGGGTG			GCACTCCGTC	
25		CCGCGGCGTA			TGGCGGACGC	GACGGAGTTG	
35	70321		GCTGCGGGGG		GGGCGATGCT TCGCCGCGGT	CGCCGTCGAC	
	70381 70441	CGGAGGTCGG		GATCTGGACA GATGTGGCGG	CGTTCGAACG	GGAGTGGTCG	
	70501	TGCTCGCCGG	TTCGCCGGAC ACGGCTCGAC		CGTTCCACTC	CCGGCACGTC	
	70561	TOGACIGOTT		CTGGAGTCGC	TOGOGTTOGG	CGCGGCGCGG	
40	70501	TGTCCACGAC	GACGGGCCGG	GACGCCGCGG	ACGACCTCAT	AACGCCCGCG	CACTGGCTGC
70	70681		TOGGOOGGTG			GGAGCTGGCC	
		TCACCACGTT					
	70801	CCGGGGAGGA	CGCCGGGACC	T'ACCACGCGG	TGCTGCGCGC	CCGGACCGGT	GAGGAGACCG
		CGGCGCTGAC					
45		TACIGGCCGG					
		GGCTGGCCCC					
		AGTCCGAGCC					
		TCGGCGTCAC					
	71161	ACTCACTGGC	GGTGCAGCGG	CTGCGCAACC	TOTEGETE	GGCAACCGGG	CTGGACCTGC
50	71221	CGGCGGCCGT	CCTGTTCGAC	CACGACACCC	CGGCCGCGCT	CACCGCGTTC	CTCCAGGACC
	71281	GGATCGAGGC	CGGCCAGGAC	CGGATCGAGG	CCGGCGAGGA	CGACGACGCG	CCCACCGTGC
		TCTCGCTCCT					
		CGGAGCGTGC					
		GATGAGCACC					
55	71521	GGACGGTCAC	CGCGCCATCC	TGGAGAGCGG	CACGGTGGGT	TCGTTCGACC	TGTTCGGCGT
	71581	CAAGCACTGG	CTGGTCGCCG	CCGCCGAGGA	CGTCAAGCTG	GTCACCAACG	ATCCGCGGTT
	71641	CAGCTCGGCC	GCGCCGTCCG	AGATGCTGCC	CGACCGGCGG	CCCGGCTGGT	TCTCCGGGAT
	71701	GGACTCACCG	GAGCACAACC	GCTACCGGCA	GAAGATCGCG	GGGGACTTCA	CAUTGUGUGU
60	71761	GGCGCGCAAG	CGGGAGGACT	TCGTCGCCGA	GGCCGCCGAC	GUUTGUUTGG	ACGACATCGA
60	/1821	GGCCGCGGGA	CCCGGCACCG	ACCTUATOCC	CGGGT ACGCC	AAGCGGCTGC	CACCCATCCC
	/1881	CATCAACGCG	CTGTACGGGC	TUACCCCTGA	OCEAN CARCE	- GIGCIGGAGG	ACTTOTALCO
	/1941	CGACATCACC GCACGCGCTG	GGCTCGGCCG	ATCTGGACAG	TOTCAMBAGE	CIGACCGACG	MCTICITOGG
	72001	GCACGCGCTG GCTGGCCTCG	COCCACCACA	COCACAMEMO	COMORGOOF	GACGAGGACC	775750700000 7757507000000
	12061	. GUTGGUCTCG	GUUGAUGAUG	GCGAGATCIC	ひに 1 くみじじらおし	JACOMOGCOM	110100000

	20101			CCChcchcm	0000000000		
			CTGTTCGCCG				
			AGCCACCCCG				
	3000		GAGGAGATGC				
5	2002		GACGTOGATG				
_,		GOTOTACTOG				CAGCCCGACA	
	72421		CTGGAGGGCA				
	72481		CGGGTGCTCA				
	72541		GCCGGCGACG				
10	72601		ACCTGGGGGG				
10	72661	GGGACGACGG	TCGCGCACAT				
			GCTACCTGCG				
			ACATCGGCAT				
	72841		TCGAGCCCGC				
1.5	72901		CGGGCCAGGC				
15			ATCCCGACGC				
			TGCGCACGCT				
			AACTGCCCGA				
			CGGAGCGCGG				
20			AGGTCTTCGC				
20			TOCACGACAT				
	73321		CCGTGGTCGC				
	73351		GGCGGGTGGC				
			GGACGGCGGC				
			CCAGCTTGCG				
25			GGCTGGCGAT				
			CCTCGGTCAG				
			AGGACTCCCC				
			GTGCGCGGCG				
			CCATGTCGGC				
30			CGGCCTCGTC				
			GUGTCATCAC				
	73981	ATGAGCCTCA	GCCCCTCGTC	ACGGCCGCGG	CCGAGCAGCA	GAAGCGCTTC	GGCGGCGTCG
	74041	ACCCGCCACA	GGGCCAGGCC	CGGCACGTCG	ACGGACCAGC	GTCGCATCCG	CTCCCCGCAG
			CGTTGTACGC				
35	74161	GCCCAGACCA	TGTGCAGTCC	GAAGAGGCTG	TCGGAGGTCT	CCTCCGGCAA	CGGCTCGGCG
	74221	AGCCACCGCT	CCGCCCGGTC	CAGGTCGCCC	AGTCGGATCG	CGGCGGCCAC	GGTGCTGCTC
	74281	AGCGGCAATG	CGGCGGCCAT	CCCCCAGGAG	GGCACGACCC	GGGGGGCGAG	CGCGGCCTCG
			CGGCGGCGGT				
	74401	GCGTGGACCG	CCTCGTCGGC	CGGGGTCCGC	ATGTTGTCGT	CACCGGCCAG	CTTGTCGACC
40	74461	CAGGACTGGA	CGGCATCGGT	GTCCTCGGCG	TAGAGCAGGG	CCAGCAACGC	CATCATGGTC
	74521	GTGGTCCGGT	CCGTCGTGAC	CCGGGGAGTGC	TGGAGCACGT	ACTCGGCTTT	GGCCTCGGCC
	74581	TGTTCGGACC	AGCCGCGCAG	CGCGTTGCTC	AGGGCCTTGT	CGGCGACGGC	GCGGTGCCGG
	74641	ACGGCTCCGG	AAAACGAGGC	GACCTCGTCC	TOGGCOGGCG	GATCGGCCGG	ACGCGGCGGA
	74701	TOGGCCGCGC	CGGGATAGAT	CAGCGCGAGG	GACAGGTCCG	CGACGCGCAG	GTGCGCCCGG
45	74761	CCCTGCTCGC	TOGGGGGGGC	GGAGCGCTGG	GCCGCCAGGA	CCTCGGCGGC	CTCGCCCGGC
	74821	CGCCCGTCCA	TOGOCAGOCA	GCAGGCGAGC	GACACGGCGT	GCTCGCTGGA	GAGGAGCCGT
	74881	TOCCGCGACG	CGGTGAGCAG	CTCGGGCACA	TGCCGGCCGG	ATCTGGCGGG	ATCGCAGAGC
	74941	CGCTCGATGG	CGGCGGTGTC	GACGCGCAGT	GCGGCGTGGA	CGGCGGGGTC	GTCGGAGGCC
	75001	CGGTAGGCGA	ACTCCAGGTA	GGTGACGGCC	TCGTCGAGCT	CGCCGCGCAG	GTGGTGCTCG
50	75061	CGCGCGGCGT	CGGTGAACAG	CCCGGCGACC	TCGGCGCCGT	GCACCCGGCC	GGTACCCATC
			CGAGCACCTT				
			CACGCCGCTC				
			ACCGCCCTTC				
			CGGTGTCGAG				
55			CGGAAGCTCG				
			CGAGCCGGTA				
			CCCGGATGTC				
			CCTGGGCCAC				
			TCTGCGCCTC				
60			CCGCCCGGAA				
2.0			CACGGGCCCG				
			CGTGGTGCAC				
			TGCGGGTGAG				
			CCGTCAGCCG				
	, 5 , 0 1	1 COCH JONIO	200102000	5,100.100100	2210100000	20001.0010	

	75961	AGGAGCTGGC	CGAGCATGCC	GTACGGCAGG	GCCCGCTCCT	CCATGGAGCA	CACCGCGCGA
	76021	AGGGTGACGA	AGCCGGCCTT	9900909993	GCGTCGAGGA	STTCGGTCTT	GCCGCAGGCG
	76081	ATOSGCCCGG	TGACGGCGCC	6806806000	03000000000	CCGCTCGGGT	GAGCGCCCGG
	~6141	TBGAGGGAAC	CGAACTCGTO	AT000000000	ATCAGGTCTG	GGGGAGATAA	GCGCGCTATC
5	7.5261	ACGAATGGAA	OTACCTCGCG	A009700T00	AAACCCATAG	GCATCACATG	GCTTGTTGAT
	16261	CTGTACGGCT	GTGATTCAGC	07330033AT	SCTSTBOTAC	AGATGGGAAG	ATGTGATCTA
	76311	3669037600	GTTCCCTCAG	GRGCCGACCG	ddddddggdgn	CACCEGEEGT	ACCCCCTGGG
	76381	CCACCAGCTC	GGCGACCCGC	TOOTGGTGGT	OGACGAGGTA	GAAGTGCCCG	CCGGGGAAGA
	76441	CCTCCACCGT	GGTCGGCGCG	GTCGTGTGCC	CGGCCCAGGC	GTGGGCCTGC	TCCACCGTCG
10	76501	TOTTOGGATO	GTCGTCACCG	ATGCACACCG	TGATCGGCGT	CTCCAGCGGC	GGCGCGGGCT
	76561	CCCACCGGTA	CGTCTCCGCC	GCGTAGTAGT	COGCCCGCAA	CGGCGCCAGG	ATCAGCGCGC
	76621	GCATTTCSTC	GTCCGCCATC	ACATOGGOGO	TOGTCCCGCC	GAGGCCGATG	ACCGCCGCCA
	76681	GCAGCTCGTC	GTCGGACGCG	AGGTGGTCCT	GGTCGGCGCG	CGGCTGCGAC	GGCGCCCGCC
	76741	GGCCCGAGAC	GATCAGGTGC	GCCACCGGGA	GCCGCTGGGG	CAGCTCGAAC	GCGAGTGTCG
15	76801	CGCCCATGCT	GTGGCCGAAC	AGCACCAGCG	GACGGTCCAG	CCCCGGCTTC	AACGCCTCGG
	76861	CCACGAGGCC	GGCGAGAACA	CGCAGGTCGC	GCACCGCCTC	OTCGTCGCGG	CGGTCCTGGC
	76921	GGCCGGGGTA	CTGCACGGCG	TACACGTCCG	CCACC 30360	GAGCGCACGG	GUUAGCGGAA
	76981	GGTAGAACGT	CGCCGATCCG	coggar" ya	GCAGCAGCAC	CACCCGTACC	GGGGCCTCGG
	77041	GCGTGGGGAA	GAACTGCCGC	AGCCAGAGTT	CCGAGCTCAC	CGCACCCCCT	CGGCCGCGAC
20	77101	CTGGGGAGCC	CGGAACCGGG	TGATCTCGGC	CAAGTGCTTC	TOCOGCATOT	CCGGGTCGGT
	77161	CACGCCCCAT	CCCTCCTCCG	GCGCCAGACA	GAGGACGCCG	ACTTTGCCGT	TGTGCACATT
	77221	GCGATGCACA	TCGCGCACCG	COGRECESAC	GTCCTCCAGC	GGGTAGGTCA	CCGACAGCGT
	77231	CGGGTGCACC	ATCCCCTTGC	AGATCAGGCG	GTTCGCCTCC	CACGCCTCAC	GATAGTTCGC
	77341	GAAGTGGGTA	CCGATGATCC	GCTTCACGGA	CATCCACAGG	TACCGATTGT	CAAAGGCGTG
25	77401	STOGTATOCO	GAGGTTGACG	CGCAGGTGAC	GATOGTGCCA	CCCCGACGTG	TCACGTIGAC
	77461	ACTCGCGCCG	AACGTCGCGC	GCCCCGGGTG	CTCGAACACG	ATGTCGGGAT	CGTCACCGCC
	77521	GGTCAGCTCC	CGGATC				

Those of skill in the art will recognize that, due to the degenerate nature of the genetic code, a variety of DNA compounds differing in their nucleotide sequences can be used to encode a given amino acid sequence of the invention. The native DNA sequence encoding the FK-520 PKS of *Streptomyces hygroscopicus* is shown herein merely to illustrate a preferred embodiment of the invention, and the present invention includes DNA compounds of any sequence that encode the amino acid sequences of the polypeptides and proteins of the invention. In similar fashion, a polypeptide can typically tolerate one or more amino acid substitutions, deletions, and insertions in its amino acid sequence without loss or significant loss of a desired activity. The present invention includes such polypeptides with alternate amino acid sequences, and the amino acid sequences shown merely illustrate preferred embodiments of the invention.

The recombinant nucleic acids, proteins, and peptides of the invention are many and diverse. To facilitate an understanding of the invention and the diverse compounds and methods provided thereby, the following general description of the FK-520 PKS genes and modules of the PKS proteins encoded thereby is provided. This general description is followed by a more detailed description of the various domains and modules of the FK-520 PKS contained in and encoded by the compounds of the invention. In this description, reference to a heterologous PKS refers to any PKS other than the FK-520 PKS. Unless otherwise indicated, reference to a PKS includes reference

to a portion of a PKS. Moreover, reference to a domain, module, or PKS includes reference to the nucleic acids encoding the same and vice-versa, because the methods and reagents of the invention provide or enable one to prepare proteins and the nucleic acids that encode them.

5

10

15

20

25

30

The FK-520 PKS is composed of three proteins encoded by three genes designated fkbA, fkbB, and fkbC. The fkbA ORF encodes extender modules 7 - 10 of the PKS. The fkbB ORF encodes the loading module (the CoA ligase) and extender modules 1 - 4 of the PKS. The fkbC ORF encodes extender modules 5 - 6 of the PKS. The fkbP ORF encodes the NRPS that attaches the pipecolic acid and cyclizes the FK-520 polyketide.

The loading module of the FK-520 PKS includes a CoA ligase, an ER domain, and an ACP domain. The starter building block or unit for FK-520 is believed to be a dihvdroxycyclohexene carboxvlic acid, which is derived from shikimate. The recombinant DNA compounds of the invention that encode the loading module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of methods and in a variety of compounds. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 loading module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for the loading module of the heterologous PKS is replaced by the coding sequence for the FK-520 loading module, provides a novel PKS coding sequence. Examples of heterologous PKS coding sequences include the rapamycin, FK-506, rifamycin, and avermeetin PKS coding sequences. In another embodiment, a DNA compound comprising a sequence that encodes the FK-520 loading module is inserted into a DNA compound that comprises the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the loading module coding sequence is utilized in conjunction with a heterologous coding sequence. In this embodiment, the invention provides, for example, either replacing the CoA ligase with a different CoA ligase, deleting the ER, or replacing the ER with a different ER. In addition, or alternatively, the ACP can be replaced by another ACP. In similar fashion, the corresponding domains in another loading or extender module can be replaced by one or more domains of the FK-520 PKS. The resulting heterologous loading module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide.

The first extender module of the FK-520 PKS includes a KS domain, an AT domain specific for methylmalonyl CoA. a DH domain, a KR domain, and an ACP domain. The recombinant DNA compounds of the invention that encode the first extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 first extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the first extender module of the FK-520 PKS or the latter is merely according sequences for modules of the mechanism provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the first extender module of the FK-520 PKS is inserted into a DNA compound that comprises the remainder of the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

5

10

15

20

25

30

35

In another embodiment, all or only a portion of the first extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2hydroxymalonyl CoA specific AT; deleting either the DH or KR or both; replacing the DH or KR or both with another DH or KR; and/or inserting an ER. In replacing or inserting KR, DH, and ER domains, it is often beneficial to replace the existing KR, DH, and ER domains with the complete set of domains desired from another module. Thus, if one desires to insert an ER domain, one may simply replace the existing KR and DH domains with a KR, DH, and ER set of domains from a module containing such domains. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a gene for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous first extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the first extender module of the FK-520 PKS.

In an illustrative embodiment of this aspect of the invention, the invention provides recombinant PKSs and recombinant DNA compounds and vectors that encode

such PKSs in which the KS domain of the first extender module has been inactivated. Such constructs are especially useful when placed in translational reading frame with the remaining modules and domains of an FK-520 or FK-520 derivative PKS. The utility of these constructs is that host cells expressing, or cell free extracts containing, the PKS encoded thereby can be fed or supplied with N-acyleysteamine thioesters of novel precursor molecules to prepare FK-520 derivatives. See U.S. patent application Serial No. 60/117,384, filed 27 Jan. 1999, and PCT patent publication Nos. US97/02358 and US99/03986, each of which is incorporated herein by reference.

The second extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the second extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 second extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the second extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the second extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, all or a portion of the second extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting the KR and/or the inactive DH; replacing the KR with another KR; and/or inserting an active DH or an active DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous second extender module coding sequence can be utilized in conjunction with a coding sequence from a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding

domains in a module of a heterologous PKS can be replaced by one or more domains of the second extender module of the FK-520 PKS.

The third extender module of the FK-520 PKS includes a KS, an AT specific for malonyl CoA, a KR, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the third extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 third extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the third extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the third extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, all or a portion of the third extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the malonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting the KR and/or the inactive DH; replacing the KR with another KR; and/or inserting an active DH or an active DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous third extender module coding sequence can be utilized in conjunction with a coding sequence from a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the third extender module of the FK-520 PKS.

The fourth extender module of the FK-520 PKS includes a KS, an AT that binds ethylmalonyl CoA, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the fourth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In

one embodiment, a DNA compound comprising a sequence that encodes the FK-520 fourth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the fourth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the fourth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the remainder of the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the fourth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the ethylmalonyl CoA specific AT with a malonyl CoA, methylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; and/or deleting the inactive DH, inserting a KR, a KR and an active DH, or a KR, an active DH, and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, a PKS for a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous fourth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the fourth extender module of the FK-520 PKS.

As illustrative examples, the present invention provides recombinant genes, vectors, and host cells that result from the conversion of the FK-506 PKS to an FK-520 PKS and vice-versa. In one embodiment, the invention provides a recombinant set of FK-506 PKS genes but in which the coding sequences for the fourth extender module or at least those for the AT domain in the fourth extender module have been replaced by those for the AT domain of the fourth extender module of the FK-520 PKS. This recombinant PKS can be used to produce FK-520 in recombinant host cells. In another embodiment, the invention provides a recombinant set of FK-520 PKS genes but in which the coding sequences for the fourth extender module or at least those for the AT domain in the fourth extender module have been replaced by those for the AT domain of

the fourth extender module of the FK-506 PKS. This recombinant PKS can be used to produce FK-506 in recombinant host cells.

Other examples of hybrid PKS enzymes of the invention include those in which the AT domain of module 4 has been replaced with a malonyl specific AT domain to provide a PKS that produces 21-desethyl-FK520 or with a methylmalonyl specific AT domain to provide a PKS that produces 21-desethyl-21-methyl-FK520. Another hybrid PKS of the invention is prepared by replacing the AT and inactive KR domain of FK-520 extender module 4 with a methylmalonyl specific AT and an active KR domain, such as, for example, from module 2 of the DEBS or oleandolide PKS enzymes, to produce 21-desethyl 21-methyl-22-desoxo-22-hydroxy-FK520. The compounds produced by these hybrid PKS enzymes are neurotrophins.

10

15

20

25

30

35

The fifth extender module of the FK-520 PKS includes a KS, an AT that binds methylmalonyl CoA, a DH, a KR, and an ACP. The recombinant DNA compounds of the invention that encode the fifth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 fifth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the fifth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS. In another embodiment, a DNA compound comprising a sequence that encodes the fifth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the fifth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting any one or both of the DH and KR; replacing any one or both of the DH and KR with either a KR and/or DH; and/or inserting an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous fifth extender module coding sequence

can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the fifth extender module of the FK-520 PKS.

5

10

15

20

25

30

35

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the DH domain of the fifth extender module have been deleted or mutated to render the DH non-functional. In one such mutated gene, the KR and DH coding sequences are replaced with those encoding only a KR domain from another PKS gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-500 malog that lacks the C-19 to C 20 double bond of FK-520 and has a C-20 hydroxyl group. Such analogs are preferred neurotrophins, because they have little or no immunosuppressant activity. This recombinant fifth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment. the present invention provides a recombinant FK-520 PKS that contains both this fifth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (lacking the C-19 to C-20 double bond of FK-506 and having a C-20 hydroxyl group) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the DH domain of module 5 has been deleted or otherwise rendered inactive and thus produces this novel polyketide.

The sixth extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds of the invention that encode the sixth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 sixth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the sixth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the sixth

extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the sixth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2hydroxymalonyl CoA specific AT; deleting any one, two, or all three of the KR, DH, and ER; and/or replacing any one, two, or all three of the KR. DH, and ER with another KR, 10 DH, and ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous sixth extender module coding 15 sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the sixth extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the DH and ER domains of the sixth extender module have been deleted or mutated to render them non-functional. In one such mutated gene, the KR, ER, and DH coding sequences are replaced with those encoding only a KR domain from another PKS gene. This can also be accomplished by simply replacing the coding sequences for extender module six with those for an extender module having a methylmalonyl specific AT and only a KR domain from a heterologous PKS gene, such as, for example, the coding sequences for extender module two encoded by the ervAl gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that has a C-18 hydroxyl group. Such analogs are preferred neurotrophins, because they have little or no immunosuppressant activity. This recombinant sixth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this sixth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-

20

25

30

35

506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (having a C-18 hydroxyl group) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the DH and ER domains of module 6 have been deleted or otherwise rendered inactive and thus produces this novel polyketide.

The seventh extender module of the FK-520 PKS includes a KS, an AT specific for 2-hydroxymalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds of the invention that encode the seventh extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA—ampound comprising a sequence that encodes the FK-520 seventh extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the seventh extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the seventh extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion or all of the seventh extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the 2-hydroxymalonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or malonyl CoA specific AT; deleting the KR, the DH, and/or the ER; and/or replacing the KR, DH, and/or ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous seventh extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the seventh extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the AT domain of the seventh extender module has been replaced with those encoding an AT domain for malonyl, methylmalonyl, or ethylmalonyl CoA from another PKS gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-15 methoxy group, having instead a hydrogen, methyl, or ethyl group at that position. respectively. Such analogs are preferred, because they are more slowly metabolized than FK-520. This recombinant seventh extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this seventh extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (C-15desmethoxy) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the AT domain of module 7 has been replaced and thus produces this novel polyketide.

5

10

15

20

25

30

35

In another illustrative embodiment, the present invention provides a hybrid PKS in which the AT and KR domains of module 7 of the FK-520 PKS are replaced by a methylmalonyl specific AT domain and an inactive KR domain, such as, for example, the AT and KR domains of extender module 6 of the rapamycin PKS. The resulting hybrid PKS produces 15-desmethoxy-15-methyl-16-oxo-FK-520, a neurotrophin compound.

The eighth extender module of the FK-520 PKS includes a KS, an AT specific for 2-hydroxymalonyl CoA, a KR, and an ACP. The recombinant DNA compounds of the invention that encode the eighth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 eighth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the eighth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the eighth extender

module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

5

10

15

20

25

30

35

In another embodiment, a portion of the eighth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the 2-hydroxymalonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or malonyl CoA specific AT; deleting or replacing the KR; and/or inserting a DH or a DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous eighth extender module coding sequence can be utilized in conjunction with a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the eighth extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the AT domain of the eighth extender module has been replaced with those encoding an AT domain for malonyl, methylmalonyl, or ethylmalonyl CoA from another PKS gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-13 methoxy group, having instead a hydrogen, methyl, or ethyl group at that position, respectively. Such analogs are preferred, because they are more slowly metabolized than FK-520. This recombinant eighth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this eighth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (C-13desmethoxy) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the AT domain of module 8 has been replaced and thus produces this novel polyketide.

The ninth extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds of the invention that encode the ninth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 ninth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the ninth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKC coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the ninth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the ninth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting any one, two, or all three of the KR, DH, and ER; and/or replacing any one, two, or all three of the KR, DH, and ER with another KR, DH, and/or ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous ninth extender module coding sequence can be utilized in conjunction with a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the ninth extender module of the FK-520 PKS.

The tenth extender module of the FK-520 PKS includes a KS, an AT specific for malonyl CoA, and an ACP. The recombinant DNA compounds of the invention that encode the tenth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 tenth extender module is inserted into a DNA compound that comprises the coding sequence

for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the tenth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the tenth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

5

ìù

15

20

25

30

35

In another embodiment, a portion or all of the tenth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the malonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; and/or inserting a KR, a KR and DH, or a KR, DH, and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous tenth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the tenth extender module of the FK-520 PKS.

The FK-520 polyketide precursor produced by the action of the tenth extender module of the PKS is then attached to pipecolic acid and cyclized to form FK-520. The enzyme FkbP is the NRPS like enzyme that catalyzes these reactions. FkbP also includes a thioesterase activity that cleaves the nascent FK-520 polyketide from the NRPS. The present invention provides recombinant DNA compounds that encode the *fkbP* gene and so provides recombinant methods for expressing the *fkbP* gene product in recombinant host cells. The recombinant *fkbP* genes of the invention include those in which the coding sequence for the adenylation domain has been mutated or replaced with coding sequences from other NRPS like enzymes so that the resulting recombinant FkbP incorporates a moiety other than pipecolic acid. For the construction of host cells that do not naturally produce pipecolic acid, the present invention provides recombinant DNA compounds that express the enzymes that catalyze at least some of the biosynthesis of pipecolic acid (see Nielsen *et al.*, 1991, *Biochem. 30*: 5789-96). The *fkbL* gene encodes a

homolog of RapL, a lysine cyclodeaminase responsible in part for producing the pipecolate unit added to the end of the polyketide chain. The *fkbB* and *fkbL* recombinant genes of the invention can be used in heterologous hosts to produce compounds such as FK-520 or, in conjunction with other PKS or NRPS genes, to produce known or novel polyketides and non-ribosmal peptides.

5

10

15

20

25

30

The present invention also provides recombinant DNA compounds that encode the P450 oxidase and methyltransferase genes involved in the biosynthesis of FK-520. Figure 2 shows the various sites on the FK-520 polyketide core structure at which these enzymes act. By providing these genes in recombinant form, the present invention provides recombinant host cells that can produce FK-520. This is accomplished by introducing the recombinant PKS, P450 oxidase, and methyltransferase genes into a heterologous host cell. In a preferred embodiment, the heterologous host cell is *Streptomyces coelicolor* CH999 or *Streptomyces lividans* K4-114, as described in U.S. Patent No. 5,830,750 and U.S. patent application Serial Nos. 08/828,898, filed 31 Mar. 1997, and 09/181,833, filed 28 Oct. 1998, each of which is incorporated herein by reference. In addition, by providing recombinant host cells that express only a subset of these genes, the present invention provides methods for making FK-520 precursor compounds not readily obtainable by other means.

In a related aspect, the present invention provides recombinant DNA compounds and vectors that are useful in generating, by homologous recombination, recombinant host cells that produce FK-520 precursor compounds. In this aspect of the invention, a native host cell that produces FK-520 is transformed with a vector (such as an SCP2* derived vector for *Streptomyces* host cells) that encodes one or more disrupted genes (i.e., a hydroxylase, a methyltransferase, or both) or merely flanking regions from those genes. When the vector integrates by homologous recombination, the native, functional gene is deleted or replaced by the non-functional recombinant gene, and the resulting host cell thus produces an FK-520 precursor. Such host cells can also be complemented by introduction of a modified form of the deleted or mutated non-functional gene to produce a novel compound.

In one important embodiment, the present invention provides a hybrid PKS and the corresponding recombinant DNA compounds that encode those hybrid PKS enzymes. For purposes of the present invention a hybrid PKS is a recombinant PKS that comprises all or part of one or more modules and thioesterase/cyclase domain of a first PKS and all or part of one or more modules, loading module, and thioesterase/cyclase

domain of a second PKS. In one preferred embodiment, the first PKS is all or part of the FK-520 PKS, and the second PKS is only a portion or all of a non-FK-520 PKS.

One example of the preferred embodiment is an FK-520 PKS in which the AT domain of module 8, which specifies a hydroxymalonyl CoA and from which the C-13 methoxy group of FK-520 is derived, is replaced by an AT domain that specifies a malonyl, methylmalonyl, or ethylmalonyl CoA. Examples of such replacement AT domains include the AT domains from modules 3, 12, and 13 of the rapaymycin PKS and from modules 1 and 2 of the erythromycin PKS. Such replacements, conducted at the level of the gene for the PKS, are illustrated in the examples below. Another illustrative example of such a hybrid PKS includes an FK-520 PKS in which the natural loading module has been replaced with a loading module of another PKS. Another example of such a hybrid PKS is an FK-520 PKS in which the AT domain of module three is replaced with an AT domain that binds methylmalonyl CoA.

5

1()

15

20

25

30

35

In another preferred embodiment, the first PKS is most but not all of a non-FK-520 PKS, and the second PKS is only a portion or all of the FK-520 PKS. An illustrative example of such a hybrid PKS includes an erythromycin PKS in which an AT specific for methylmalonyl CoA is replaced with an AT from the FK-520 PKS specfic for malonyl CoA.

Those of skill in the art will recognize that all or part of either the first or second PKS in a hybrid PKS of the invention need not be isolated from a naturally occurring source. For example, only a small portion of an AT domain determines its specificity. See U.S. provisional patent application Serial No. 60/091,526, incorporated herein by reference. The state of the art in DNA synthesis allows the artisan to construct *de novo* DNA compounds of size sufficient to construct a useful portion of a PKS module or domain. For purposes of the present invention, such synthetic DNA compounds are deemed to be a portion of a PKS.

Thus, the hybrid modules of the invention are incorporated into a PKS to provide a hybrid PKS of the invention. A hybrid PKS of the invention can result not only:

- (i) from fusions of heterologous domain (where heterologous means the domains in that module are from at least two different naturally occurring modules) coding sequences to produce a hybrid module coding sequence contained in a PKS gene whose product is incorporated into a PKS, but also:
 - (ii) from fusions of heterologous module (where heterologous module means two modules are adjacent to one another that are not adjacent to one another in naturally

occurring PKS enzymes) coding sequences to produce a hybrid coding sequence contained in a PKS gene whose product is incorporated into a PKS,

(iii) from expression of one or more FK-520 PKS genes with one or more non-FK-520 PKS genes, including both naturally occurring and recombinant non-FK-520 PKS genes, and

(iv) from combinations of the foregoing.

5

10

15

20

25

30

Various hybrid PKSs of the invention illustrating these various alternatives are described herein.

Examples of the production of a hybrid PKS by co-expression of PKS genes from the FK-520 PKS and another non-FK-520 T IS include hybrid PKS enzymes produced by coexpression of FK-520 and rapamycin PKS genes. Preferably, such hybrid PKS enzymes are produced in recombinant *Streptomyces* host cells that produce FK-520 or FK-506 but have been mutated to inactivate the gene whose function is to be replaced by the rapamycin PKS gene introduced to produce the hybrid PKS. Particular examples include (i) replacement of the *fkbC* gene with the *rapB* gene; and (ii) replacement of the *fkbA* gene with the *rapC* gene. The latter hybrid PKS produces 13,15-didesmethoxy-FK-520, if the host cell is an FK-520 producing host cell, and 13,15-didesmethoxy-FK-506, if the host cell is an FK-506 producing host cell. The compounds produced by these hybrid PKS enzymes are immunosuppressants and neurotrophins but can be readily modified to act only as neurotrophins, as described in Example 6, below.

Other illustrative hybrid PKS enzymes of the invention are prepared by replacing the *fkbA* gene of an FK-520 or FK-506 producing host cell with a hybrid *fkbA* gene in which: (a) the extender module 8 through 10, inclusive, coding sequences have been replaced by the coding sequences for extender modules 12 to 14, inclusive, of the rapamycin PKS; and (b) the module 8 coding sequences have been replaced by the module 8 coding sequence of the rifamycin PKS. When expressed with the other, naturally occurring FK-520 or FK-506 PKS genes and the genes of the modification enzymes, the resulting hybrid PKS enzymes produce, respectively, (a) 13-desmethoxy-FK-520 or 13-desmethoxy-FK-506; and (b) 13-desmethoxy-13-methyl-FK-520 or 13-desmethoxy-13-methyl-FK-506. In a preferred embodiment, these recombinant PKS genes of the invention are introduced into the producing host cell by a vector such as pHU204, which is a plamsid pRM5 derivative that has the well-characterized SCP2* replicon, the *colE1* replicon, the *tsr* and *bla* resistance genes, and a *cos* site. This vector can be used to introduce the recombinant *fkbA* replacement gene in an FK-520 or FK-506 producing host cell derived therefore in which the endeageness fith A sequences fith A producing host cell (or a host cell derived therefore in which the endeageness fith A

gene has either been rendered inactive by mutation, deletion or homologous recombination with the gene that replaces it) to produce the desired hybrid PKS.

In constructing hybrid PKSs of the invention, certain general methods may be helpful. For example, it is often beneficial to retain the framework of the module to be altered to make the hybrid PKS. Thus, if one desires to add DH and ER functionalities to a module, it is often preferred to replace the KR domain of the original module with a KR, DH, and ER domain-containing segment from another module, instead of merely inserting DH and ER domains. One can alter the stereochemical specificity of a module by replacement of the KS domain with a KS domain from a module that specifies a different stereochemistry. See Lau et at., 1... Dissecting the role of acyltransferase domains of modular polyketide synthases in the choice and stereochemical fate of extender units," Biochemistry 38(5):1643-1651, incorporated herein by reference. Stereochemistry can also be changed by changing the KR domain. Also, one can alter the specificity of an AT domain by changing only a small segment of the domain. See Lau et al., supra. One can also take advantage of known linker regions in PKS proteins to link modules from two different PKSs to create a hybrid PKS. See Gokhale et al., 16 Apr. 1999, "Dissecting and Exploiting Intermodular Communication in Polyketide Synthases," Science 284: 482-485, incorporated herein by reference.

The following Table lists references describing illustrative PKS genes and corresponding enzymes that can be utilized in the construction of the recombinant PKSs and the corresponding DNA compounds that encode them of the invention. Also presented are various references describing tailoring enzymes and corresponding genes that can be employed in accordance with the methods of the present invention.

Avermectin

5

10

15

20

25

30

U.S. Pat. No. 5,252,474 to Merck.

MacNeil et al., 1993, Industrial Microorganisms: Basic and Applied Molecular Genetics, Baltz, Hegeman, & Skatrud, eds. (ASM), pp. 245-256, A Comparison of the Genes Encoding the Polyketide Synthases for Avermectin, Erythromycin, and Nemadectin.

MacNeil et al., 1992, Gene 115: 119-125, Complex Organization of the Streptomyces avermitilis genes encoding the avermectin polyketide synthase.

Ikeda *et al.*, Aug. 1999, Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*, *Proc. Natl. Acad. Sci. USA 96*: 9509-9514.

35 Candicidin (FR008)

Hu et al., 1994, Mol. Microbiol. 14: 163-172.

Epothilone

U.S. Pat. App. Serial No. 60'130,560, filed 22 April 1999.

Erythromycin

5 PCT Pub. No. 93/13663 to Abbott.

US Pat. No. 5,824,513 to Abbott.

Donadio et al., 1991, Science 252:675-9.

Cortes et al., 8 Nov. 1990, Nature 348:176-8, An unusually large multifunctional polypeptide in the erythromycin producing polyketide synthase of Saccharopolyspora erythraea.

Glycosylation Enzymes

PCT Pat. App. Pub. No. 97/23630 to Abbott.

FK-506

10

20

Motamedi *et al.*, 1998, The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK-506, *Eur. J. biochem.* 256: 528-534.

Motamedi *et al.*, 1997, Structural organization of a multifunctional polyketide synthase involved in the biosynthesis of the macrolide immunosuppressant FK-506, *Eur. J. Biochem. 244*: 74-80.

Methyltransferase

US 5,264,355, issued 23 Nov. 1993, Methylating enzyme from *Streptomyces* MA6858. 31-O-desmethyl-FK-506 methyltransferase.

Motamedi *et al.*, 1996, Characterization of methyltransferase and hydroxylase genes involved in the biosynthesis of the immunosuppressants FK-506 and FK-520, *J. Bacteriol. 178*: 5243-5248.

25 Streptomyces hygroscopicus

U.S. patent application Serial No. 09/154,083, filed 16 Sep. 1998.

Lovastatin

U.S. Pat. No. 5,744,350 to Merck.

Narbomycin

30 U.S. patent application Serial No. 60/107,093, filed 5 Nov. 1998, and Serial No. 60/120,254, filed 16 Feb. 1999.

Nemadectin

MacNeil et al., 1993, supra.

Niddamycin

Kakavas et al., 1997, Identification and characterization of the niddamycin polyketide synthase genes from *Streptomyces caelestis*, *J. Bacteriol*. 179: 7515-7522.

Oleandomycin

Swan et al., 1994, Characterisation of a *Streptomyces antibioticus* gene encoding a type I polyketide synthase which has an unusual coding sequence, *Mol. Gen. Genet.* 242: 358-362.

U.S. patent application Serial No. 60/120,254, filed 16 Feb. 1999.

Olano et al., 1998, Analysis of a Streptomyces antibioticus chromosomal region involved in oleandomycin biosynthesis, which encodes two glycosyltransferases responsible for glycosylation of the macrolactone ring, Mol. Gen. Genet. 259(3): 299-308.

Picromycin

10

PCT patent application US99/15047, filed 2 Jul. 1999.

Xue et al., 1998, Hydroxylation of macrolactones YC-17 and narbomycin is mediated by the pikC-encoded cytochrome P450 in Streptomyces venezuelae, Chemistry & Biology 5(11): 661-667.

Xue et al., Oct. 1998, A gene cluster for macrolide antibiotic biosynthesis in Streptomyces venezuelae: Architecture of metabolic diversity, Proc. Natl. Acad. Sci. USA 95: 12111 12116.

20 Platenolide

EP Pat. App. Pub. No. 791,656 to Lilly.

Rapamycin

Schwecke *et al.*, Aug. 1995, The biosynthetic gene cluster for the polyketide rapamycin, *Proc. Natl. Acad. Sci. USA 92*:7839-7843.

Aparicio et al., 1996, Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of the enzymatic domains in the modular polyketide synthase, *Gene 169*: 9-16.

Rifamycin

30

August et al., 13 Feb. 1998, Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the rif biosynthetic gene cluster of Amycolatopsis mediterranei S669, Chemistry & Biology, 5(2): 69-79.

Sorangium PKS

U.S. patent application Serial No. 09/144,085, filed 31 Aug. 1998.

Soraphen

35 U.S. Pat. No. 5,716,849 to Novartis.

Schupp et al., 1995, J. Bacteriology 177: 3673-3679. A Sorangium cellulosum (Myxobacterium) Gene Cluster for the Biosynthesis of the Macrolide Antibiotic Soraphen A: Cloning, Characterization, and Homology to Polyketide Synthase Genes from Actinomycetes.

5 Spíramycin

U.S. Pat. No. 5,098,837 to Lilly.

Activator Gene

U.S. Pat. No. 5,514,544 to Lilly.

Tylosin

15

20

25

30

35

10 EP Pub. No. 791,655 to Lilly.

U.S. Pat. No. 5,876,991 to Lilly.

Kuhstoss *et al.*, 1996, *Gene 183*:231-6., Production of a novel polyketide through the construction of a hybrid polyketide synthase.

Tailoring enzymes

Merson-Davies and Cundliffe, 1994, *Mol. Microbiol. 13*: 349-355. Analysis of five tylosin biosynthetic genes from the *tylBA* region of the *Streptomyces fradiae* genome.

As the above Table illustrates, there are a wide variety of polyketide synthase genes that serve as readily available sources of DNA and sequence information for use in constructing the hybrid PKS-encoding DNA compounds of the invention. Methods for constructing hybrid PKS-encoding DNA compounds are described without reference to the FK-520 PKS in PCT patent publication No. 98/51695; U.S. Patent Nos. 5,672,491 and 5,712,146 and U.S. patent application Serial Nos. 09/073,538, filed 6 May 1998, and 09/141,908, filed 28 Aug 1998, each of which is incorporated herein by reference.

The hybrid PKS-encoding DNA compounds of the invention can be and often are hybrids of more than two PKS genes. Moreover, there are often two or more modules in the hybrid PKS in which all or part of the module is derived from a second (or third) PKS. Thus, as one illustrative example, the present invention provides a hybrid FK-520 PKS that contains the naturally occurring loading module and FkbP as well as modules one, two, four, six, seven, and eight, nine, and ten of the FK-520 PKS and further contains hybrid or heterologous modules three and five. Hybrid or heterologous module three contains an AT domain that is specific of methylmalonyl CoA and can be derived for example, from the erythromycin or rapamycin PKS genes. Hybrid or heterologous module five contains an AT domain that is specific for malonyl CoA and can be derived for example, from the picromycin or rapamycin PKS genes.

While an important embodiment of the present invention relates to hybrid PKS enzymes and corresponding genes, the present invention also provides recombinant FK-520 PKS genes in which there is no second PKS gene sequence present but which differ from the FK-520 PKS gene by one or more deletions. The deletions can encompass one or more modules and/or can be limited to a partial deletion within one or more modules. When a deletion encompasses an entire module, the resulting FK-520 derivative is at least two carbons shorter than the gene from which it was derived. When a deletion is within a module, the deletion typically encompasses a KR, DH, or ER domain, or both DH and ER domains, or both KR and DH domains, or all three KR, DH, and ER domains.

To construct a hybrid PKS or FK-520 derivative PKS gene of the invention, one can employ a technique, described in PCT Pub. No. 98/27203 and U.S. patent application Serial No. 08/989,332, filed 11 Dec. 1997, each of which is incorporated herein by reference, in which the large PKS gene is divided into two or more, typically three, segments, and each segment is placed on a separate expression vector. In this manner, each of the segments of the gene can be altered, and various altered segments can be combined in a single host cell to provide a recombinant PKS gene of the invention. This technique makes more efficient the construction of large libraries of recombinant PKS genes, vectors for expressing those genes, and host cells comprising those vectors.

Thus, in one important embodiment, the recombinant DNA compounds of the invention are expression vectors. As used herein, the term expression vector refers to any nucleic acid that can be introduced into a host cell or cell-free transcription and translation medium. An expression vector can be maintained stably or transiently in a cell, whether as part of the chromosomal or other DNA in the cell or in any cellular compartment, such as a replicating vector in the cytoplasm. An expression vector also comprises a gene that serves to produce RNA that is translated into a polypeptide in the cell or cell extract. Furthermore, expression vectors typically contain additional functional elements, such as resistance-conferring genes to act as selectable markers.

The various components of an expression vector can vary widely, depending on the intended use of the vector. In particular, the components depend on the host cell(s) in which the vector will be used or is intended to function. Vector components for expression and maintenance of vectors in *E. coli* are widely known and commercially available, as are vector components for other commonly used organisms, such as yeast cells and *Streptomyces* cells.

In a preferred embodiment, the expression vectors of the invention are used to construct recombinant *Streptomyces* host cells that express a recombinant PKS of the invention. Preferred *Streptomyces* host cells vector combinations of the invention include *S. coelicolor* CH999 and *S. lividans* K4-114 host cells, which do not produce actinorhodin, and expression vectors derived from the pRM1 and pRM5 vectors, as described in U.S. Patent No. 5,830,750 and U.S. patent application Serial Nos. 08/828,898, filed 31 Mar. 1997, and 09 181,833, filed 28 Oct. 1998, each of which is incorporated herein by reference.

5

30

35

The present invention provides a wide variety of expression vectors for use in 10 Streptom: 3. For replicating vectors, the origin of replication can be, for example and without limitation, a low copy number vector, such as SCP2* (see Hopwood et al., Genetic Manipulation of Streptomyces: A Laboratory manual (The John Innes Foundation, Norwich, U.K., 1985); Lydiate et al., 1985, Gene 35: 223-235; and Kieser and Melton, 1988, Gene 65: 83-91, each of which is incorporated herein by reference), 15 SLP1.2 (Thompson et al., 1982, Gene 20: 51-62, incorporated herein by reference), and SG5(ts) (Muth et al., 1989, Mol. Gen. Genet. 219: 341-348, and Bierman et al., 1992, Gene 116: 43-49, each of which is incorporated herein by reference). or a high copy number vector, such as pIJ101 and pJV1 (see Katz et al., 1983, J. Gen. Microbiol. 129: 2703-2714; Vara et al., 1989, J. Bacteriol. 171: 5782-5781; and Servin-Gonzalez, 1993, Plasmid 30: 131-140, each of which is incorporated herein by reference). Generally, 20 however, high copy number vectors are not preferred for expression of genes contained on large segments of DNA. For non-replicating and integrating vectors, it is useful to include at least an E. coli origin of replication, such as from pUC, p1P, p1I, and pBR. For phage based vectors, the phages phiC31 and KC515 can be employed (see Hopwood 25 et al., supra).

Typically, the expression vector will comprise one or more marker genes by which host cells containing the vector can be identified and/or selected. Useful antibiotic resistance conferring genes for use in *Streptomyces* host cells include the *ermE* (confers resistance to erythromycin and other macrolides and lincomycin), *tsr* (confers resistance to thiostrepton), *aadA* (confers resistance to spectinomycin and streptomycin), *aacC4* (confers resistance to apramycin, kanamycin, gentamicin, geneticin (G418), and neomycin), *hyg* (confers resistance to hygromycin), and *vph* (confers resistance to viomycin) resistance conferring genes.

The recombinant PKS gene on the vector will be under the control of a promoter, typically with an attendant ribosome binding site sequence. The present invention

provides the endogenous promoters of the FK-520 PKS and related biosynthetic genes in recombinant form, and these promoters are preferred for use in the native hosts and in heterologous hosts in which the promoters function. A preferred promoter of the invention is the *fkbO* gene promoter, comprised in a sequence of about 270 bp between the start of the open reading frames of the *fkbO* and *fkbB* genes. The *fkbO* promoter is believed to be bi-directional in that it promotes transcription of the genes *fkbO*, *fkbP*, and *fkbA* in one direction and *fkbB*, *fkbC*, and *fkbL* in the other. Thus, in one aspect, the present invention provides a recombinant expression vector comprising the promoter of the *fkbO* gene of an FK-520 producing organism positioned to transcribe a gene other than *fkbO* gene of an FK-520 producing organism positioned to transcribe a gene other than *fkbO* gene and producing organism positioned to transcribe a gene other than *fkbO* gene of an FK-520 producing organism positioned to transcribe a gene other than *fkbO* gene of an FK-520 producing organism positioned to transcribe a gene other than *fkbO* gene of an FK-520 producing organism positioned to transcribe a gene other than *fkbO* gene of an FK-520 producing organism positioned to transcribe a gene other than *fkbO* gene of an FK-520 producing organism positioned to transcribe a gene other than *fkbO* gene of an FK-520 producing organism positioned to transcribe a gene other than *fkbO* gene of an FK-520 producing organism positioned to transcribe a gene other than *fkbO* gene of an FK-520 producing organism positioned to transcribe a gene of an FK-520 producing organism positioned to transcribe a gene of an FK-520 producing organism positioned to transcribe a gene of an FK-520 producing organism positioned to transcribe a gene of an FK-520 producing organism positioned to transcribe a gene of an FK-520 producing organism positioned to transcribe a gene of an FK-520 producing organism positioned to transcribe a gene of an FK-520

5

. ()

15

20

25

30

Heterologous promoters can also be employed and are preferred for use in host cells in which the endogenous FK-520 PKS gene promoters do not function or function poorly. A preferred heterologous promoter is the actl promoter and its attendant activator gene actII-ORF4, which is provided in the pRM1 and pRM5 expression vectors, supra. This promoter is activated in the stationary phase of growth when secondary metabolites are normally synthesized. Other useful Streptomyces promoters include without limitation those from the ermE gene and the melC1 gene, which act constitutively, and the tipA gene and the merA gene, which can be induced at any growth stage. In addition, the T7 RNA polymerase system has been transferred to Streptomyces and can be employed in the vectors and host cells of the invention. In this system, the coding sequence for the T7 RNA polymerase is inserted into a neutral site of the chromosome or in a vector under the control of the inducible merA promoter, and the gene of interest is placed under the control of the T7 promoter. As noted above, one or more activator genes can also be employed to enhance the activity of a promoter. Activator genes in addition to the actII-ORF4 gene discussed above include dnrI, redD, and ptpA genes (see U.S. patent application Serial No. 09/181,833, supra) to activate promoters under their control.

In addition to providing recombinant DNA compounds that encode the FK-520 PKS, the present invention also provides DNA compounds that encode the ethylmalonyl CoA and 2-hydroxymalonyl CoA utilized in the synthesis of FK-520. Thus, the present invention also provides recombinant host cells that express the genes required for the biosynthesis of ethylmalonyl CoA and 2-hydroxymalonyl CoA. Figures 3 and 4 show the

location of these genes on the cosmids of the invention and the biosynthetic pathway that produces ethylmalonyl Co.A.

For 2-hydroxymalonyl CoA biosynthesis, the *fkbH*, *fkbI*, *fkbJ*, and *fkbK* genes are sufficient to confer this ability on *Streptomcyces* host cells. For conversion of 2-

hydroxymalonyl to 2-methoxymalonyl, the *fkbG* gene is also employed. While the complete coding sequence for *fkbH* is provided on the cosmids of the invention, the sequence for this gene provided herein may be missing a T residue, based on a comparison made with a similar gene cloned from the ansamitocin gene cluster by Dr. H. Floss. Where the sequence herein shows one T, there may be two, resulting in an

10

15

20

25

30

35

extension of the *fkbH* reading frame to encode the amino acid sequence:

MTIVKCLVWDLDNTLWRGTVLEDDEVVLTDEIREVITTLDDRGILQAVASKNDH

DLAWERLERLGVAEYFVLARIGWGPKSQSVREIATELNFAPTTIAFIDDQPAERA

EVAFHLPEVRCYPAEQAATLLSLPEFSPPVSTVDSRRRRLMYQAGFARDQAREA

YSGPDEDFLRSLDLSMTIAPAGEEELSRVEELTLRTSQMNATGVHYSDADLRAL

LTDPAHEVLVVTMGDRFGPHGAVGIILLEKKPSTWHLKLLATSCRVVSFGAGAT ILNWLTDQGARAGAHLVADFRRTDRNRMMEIAYRFAGFADSDCPCVSEVAGAS AAGVERLHLEPSARPAPTTLTLTAADIAPVTVSAAG.

For ethylmalonyl CoA biosynthesis, one requires only a crotonyl CoA reductase, which can be supplied by the host cell but can also be supplied by recombinant expression of the *fkbS* gene of the present invention. To increase yield of ethylmalonyl CoA, one can also express the *fkbE* and *fkbU* genes as well. While such production can be achieved using only the recombinant genes above, one can also achieve such production by placing into the recombinant host cell a large segment of the DNA provided by the cosmids of the invention. Thus, for 2-hydroxymalonyl and 2-methoxymalonyl CoA biosynthesis, one can simply provide the cells with the segment of DNA located on the left side of the FK-520 PKS genes shown in Figure 1. For ethylmalonyl CoA biosynthesis, one can simply provide the cells with the segment of DNA located on the right side of the FK-520 PKS genes shown in Figure 1 or, alternatively, both the right and left segments of DNA.

The recombinant DNA expression vectors that encode these genes can be used to construct recombinant host cells that can make these important polyketide building blocks from cells that otherwise are unable to produce them. For example, *Streptomyces coelicolor* and *Streptomyces lividans* do not synthesisze ethylmalonyl CoA or 2-hydroxymalonyl CoA. The invention provides methods and vectors for constructing recombinant *Streptomyces coelicolor* and *Streptomyces lividans* that are able to

synthesize either or both ethylmalonyl CoA and 2-hydroxymalonyl CoA. These host cells are thus able to make polyketides, those requiring these substrates, that cannot otherwise be made in such cells.

In a preferred embodiment, the present invention provides recombinant

Streptomyces host cells, such as S. coelicolor and S. lividans, that have been transformed with a recombinant vector of the invention that codes for the expression of the ethylmalonyl CoA biosynthetic genes. The resulting host cells produce ethylmalonyl CoA and so are preferred host cells for the production of polyketides produced by PKS enzymes that comprise one or more AT domains specific for ethylmalonyl CoA.

Illustrative PKS enzymes of this type include the FK-520 PKS and a recombinant PKS in

which one or more AT domains is specific for ethylmalonyl CoA.

15

20

25

30

In a related embodiment, the present invention provides *Streptomyces* host cells in which one or more of the ethylmalonyl or 2-hydroxymalonyl biosynthetic genes have been deleted by homologous recombination or rendered inactive by mutation. For example, deletion or inactivation of the *fkbG* gene can prevent formation of the methoxyl groups at C-13 and C-15 of FK-520 (or, in the corresponding FK-506 producing cell, FK-506), leading to the production of 13,15-didesmethoxy-13,15-dihydroxy-FK-520 (or, in the corresponding FK-506 producing cell, 13,15-didesmethoxy-13,15-dihydroxy-FK-506). If the *fkbG* gene product acts on 2-hydroxymalonyl and the resulting 2-methoxymalonyl substrate is required for incorporation by the PKS, the AT domains of modules 7 and 8 may bind malonyl CoA and methylmalonyl CoA. Such incorporation results in the production of a mixture of polyketides in which the methoxy groups at C-13 and C-15 of FK-520 (or FK-506) are replaced by either hydrogen or methyl.

This possibility of non-specific binding results from the construction of a hybrid PKS of the invention in which the AT domain of module 8 of the FK-520 PKS replaced the AT domain of module 6 of DEBS. The resulting PKS produced, in *Streptomyces lividans*, 6-dEB and 2-desmethyl-6-dEB, indicating that the AT domain of module 8 of the FK-520 PKS could bind malonyl CoA and methylmalonyl CoA substrates. Thus, one could possibly also prepare the 13,15-didesmethoxy-FK-520 and corresponding FK-506 compounds of the invention by deleting or otherwise inactivating one or more or all of the genes required for 2-hydroxymalonyl CoA biosynthesis, i.e., the *fkbH*, *fkbI*, *fkbJ*, and *fkbK* genes. In any event, the deletion or inactivation of one or more biosynthetic genes required for ethylmalonyl and/or 2-hydroxymalonyl production prevents the formation of polyketides requiring ethylmalonyl and/or 2-hydroxymalonyl for biosynthesis, and the

resulting host cells are thus preferred for production of polyketides that do not require the same.

5

10

30

35

The host cells of the invention can be grown and fermented under conditions known in the art for other purposes to produce the compounds of the invention. See, e.g., U.S. Patent Nos. 5,194,378; 5,116,756; and 5,494,820, incorporated herein by reference, for suitable fermentation processes. The compounds of the invention can be isolated from the fermentation broths of these cultured cells and purified by standard procedures. Preferred compounds of the invention include the following compounds: 13-desmethoxy-FK-506; 13-desmethoxy-FK-520; 13-desmethoxy-FK-520; 13-desmethoxy-FK-506; 13-desmethoxy-18-hydroxy-FK-506; 13-desmethoxy-18-hydroxy-FK-506; and 13,15-didesmethoxy-18-hydroxy-FK-520. These compounds can be further modified as described for tacrolimus and FK-520 in U.S. Patent Nos. 5,225,403; 5,189,042; 5,164,495; 5,068,323; 4,980,466; and 4,920,218, incorporated herein by reference.

15 Other compounds of the invention are shown in Figure S. Parts A and B. In Figure 8, Part A, illustrative C-32-substituted compounds of the invention are shown in two columns under the heading R. The substituted compounds are preferred for topical administration and are applied to the dermis for treatment of conditions such as psoriasis. In Figure 8, Part B, illustrative reaction schemes for making the compounds shown in 20 Figure 8. Part A, are provided. In the upper scheme in Figure 8, Part B, the C-32 substitution is a tetrazole moiety, illustrative of the groups shown in the left column under R in Figure 8, Part A. In the lower scheme in Figure 8, Part B, the C-32 substitution is a disubstituted amino group, where R₃ and R₄ can be any group similar to the illustrative groups shown attached to the amine in the right column under R in Figure 25 8, Part A. While Figure 8 shows the C-32-substituted compounds in which the C-15methoxy is present, the invention includes these C-32-substituted compounds in which C-15 is ethyl, methyl, or hydrogen. Also, while C-21 is shown as substituted with ethyl or allyl, the compounds of the invention includes the C-32-substituted compounds in which C-21 is substituted with hydrogen or methyl.

To make these C-32-substituted compounds, Figure 8, Part B, provides illustrative reaction schemes. Thus, a selective reaction of the starting compound (see Figure 8, Part B, for an illustrative starting compound) with trifluoromethanesulfonic anhydride in the presence of a base yields the C-32 O-triflate derivative, as shown in the upper scheme of Figure 8, Part B. Displacement of the triflate with 1H-tetrazole or triazole derivatives provides the C-32 tetrazole or teiazole derivative. As shown in the

lower scheme of Figure 8, Part B, reacting the starting compound with pnitrophenylchloroformate yields the corresponding carbonate, which, upon displacement with an amino compound, provides the corresponding carbamate derivative.

5

10

15

20

25

30

35

The compounds can be readily formulated to provide the pharmaceutical compositions of the invention. The pharmaceutical compositions of the invention can be used in the form of a pharmaceutical preparation, for example, in solid, semisolid, or liquid form. This preparation contains one or more of the compounds of the invention as an active ingredient in admixture with an organic or inorganic carrier or excipient suitable for external, enteral, or parenteral application. The active ingredient may be compounded, for example, with the usual non-toxic, pharmaceutically acceptable carriers for tablets, pellets, capsules, suppositories, solutions, emulsions, suspensions, and any other form suitable for use. Suitable formulation processes and compositions for the compounds of the present invention are described with respect to tacrolimus in U.S. Patent Nos. 5,939,427; 5,922,729; 5,385,907; 5,338,684; and 5,260,301, incorporated herein by reference. Many of the compounds of the invention contain one or more chiral centers, and all of the stereoisomers are included within the scope of the invention, as pure compounds as well as mixtures of stereoisomers. Thus the compounds of the invention may be supplied as a mixture of stereoisomers in any proportion.

The carriers which can be used include water, glucose, lactose, gum acacia, gelatin, mannitol, starch paste, magnesium trisilicate, talc, corn starch, keratin, colloidal silica, potato starch, urea, and other carriers suitable for use in manufacturing preparations, in solid, semi-solid, or liquified form. In addition, auxiliary stabilizing, thickening, and coloring agents and perfumes may be used. For example, the compounds of the invention may be utilized with hydroxypropyl methylcellulose essentially as described in U.S. Patent No. 4,916,138, incorporated herein by reference, or with a surfactant essentially as described in EPO patent publication No. 428,169, incorporated herein by reference.

Oral dosage forms may be prepared essentially as described by Hondo *et al.*, 1987, *Transplantation Proceedings XIX*, Supp. 6: 17-22, incorporated herein by reference. Dosage forms for external application may be prepared essentially as described in EPO patent publication No. 423,714, incorporated herein by reference. The active compound is included in the pharmaceutical composition in an amount sufficient to produce the desired effect upon the disease process or condition.

For the treatment of conditions and diseases relating to immunosuppression or neuronal damage, a compound of the invention may be administered orally, topically,

parenterally, by inhalation spray, or rectally in dosage unit formulations containing conventional non-toxic pharmaceutically acceptable carriers, adjuvant, and vehicles. The term parenteral, as used herein, includes subcutaneous injections, and intravenous, intramuscular, and intrasternal injection or infusion techniques.

5

10

15

20

25

30

Dosage levels of the compounds of the present invention are of the order from about 0.01 mg to about 50 mg per kilogram of body weight per day, preferably from about 0.1 mg to about 10 mg per kilogram of body weight per day. The dosage levels are useful in the treatment of the above-indicated conditions (from about 0.7 mg to about 3.5 mg per patient per day, assuming a 70 kg patient). In addition, the compounds of the present invention may be administered on an intermittent basis. Fig., at semi-weekly, weekly, semi-monthly, or monthly intervals.

The amount of active ingredient that may be combined with the carrier materials to produce a single dosage form will vary depending upon the host treated and the particular mode of administration. For example, a formulation intended for oral administration to humans may contain from 0.5 mg to 5 g of active agent compounded with an appropriate and convenient amount of carrier material, which may vary from about 5 percent to about 95 percent of the total composition. Dosage unit forms will generally contain from about 0.5 mg to about 500 mg of active ingredient. For external administration, the compounds of the invention can be formulated within the range of, for example, 0.00001% to 60% by weight, preferably from 0.001% to 10% by weight, and most preferably from about 0.005% to 0.8% by weight. The compounds and compositions of the invention are useful in treating disease conditions using doses and administration schedules as described for tacrolimus in U.S. Patent Nos. 5,542,436; 5,365,948; 5,348,966; and 5,196,437, incorporated herein by reference. The compounds of the invention can be used as single therapeutic agents or in combination with other therapeutic agents. Drugs that can be usefully combined with compounds of the invention include one or more immunosuppressant agents such as rapamycin, cyclosporin A, FK-506, or one or more neurotrophic agents.

It will be understood, however, that the specific dosage level for any particular patient will depend on a variety of factors. These factors include the activity of the specific compound employed; the age, body weight, general health, sex, and diet of the subject; the time and route of administration and the rate of excretion of the drug; whether a drug combination is employed in the treatment; and the severity of the particular disease or condition for which therapy is sought.

A detailed description of the invention having been provided above, the following examples are given for the purpose of illustrating the present invention and shall not be construed as being a limitation on the scope of the invention or claims.

5 Example 1

10

15

20

25

30

35

Replacement of Methoxyl with Hydrogen or Methyl at C-13 of FK-520

The C-13 methoxyl group is introduced into FK-520 via an AT domain in extender module 8 of the PKS that is specific for hydroxymalonyl and by methylation of the hydroxyl group by an S-adenosyl methionine (SAM) dependent methyltransferase. Metabolism of FK-506 and FK-520 primarily involves oxidation at the C-13 position into an inactive derivative that is further degraded by host P450 and other enzymes. The present invention provides compounds related in structure to FK-506 and FK-520 that do not contain the C-13 methoxy group and exhibit greater stability and a longer half-life *in vivo*. These compounds are useful medicaments due to their immunosuppressive and neurotrophic activities, and the invention provides the compounds in purified form and as pharmaceutical compositions.

The present invention also provides the novel PKS enzymes that produce these novel compounds as well as the expression vectors and host cells that produce the novel PKS enzymes. The novel PKS enzymes include, among others, those that contain an AT domain specific for either malonyl CoA or methylmalonyl CoA in module 8 of the FK-506 and FK-520 PKS. This example describes the construction of recombinant DNA compounds that encode the novel FK-520 PKS enzymes and the transformation of host cells with those recombinant DNA compounds to produce the novel PKS enzymes and the polyketides produced thereby.

To construct an expression cassette for performing module 8 AT domain replacements in the FK-520 PKS, a 4.6 kb SphI fragment from the FK-520 gene cluster was cloned into plasmid pLitmus 38 (a cioning vector available from New England Biolabs). The 4.6 kb SphI fragment, which encodes the ACP domain of module 7 followed by module 8 through the KR domain, was isolated from an agarose gel after digesting the cosmid pKOS65-C31 with Sph I. The clone having the insert oriented so the single SacI site was nearest to the SpeI end of the polylinker was identified and designated as plasmid pKOS60-21-67. To generate appropriate cloning sites, two linkers were ligated sequentially as follows. First, a linker was ligated between the SpeI and SacI sites to introduce a BgIII site at the 5' end of the cassette, to eliminate interfering polylinker sites, and to reduce the total insert size to 4.5 kb (the limit of the phage

KC515). The ligation reactions contained 5 picomolar unphosphorylated linker DNA and 0.1 picomolar vector DNA, i.e., a 50-fold molar excess of linker to vector. The linker had the following sequence:

5'-CTAGTGGGCAGATCTGGCAGCT-3' 3'-ACCCGTCTAGACCG-5'

The resulting plasmid was designated pKOS60-27-1.

Next, a linker of the following sequence was ligated between the unique *Sphī* and *Aff*III sites of plasmid pKOS60-27-1 to introduce an *Nsi*I site at the 3' end of the module 8 cassette. The linker employed was:

5'-GGGATGCATGGC-3' 3'-GTACCCCTACGTACCGAATT-5'

The resulting plasmid was designated pKOS60-29-55.

To allow in-frame insertions of alternative AT domains, sites were engineered at the 5' end (Avr II or Nhe I) and 3' end (Xho I) of the AT domain using the polymerase chain reaction (PCR) as follows. Plasmid pKOS60-29-55 was used as a template for the PCR and sequence 5' to the AT domain was amplified with the primers SpeBgl-fwd and either Avr-rey or Nhe-rey:

SpeBgl-fwd 5'-CGACTCACTAGTGGGCAGATCTGG-3'

Avr-rev 5'-CACGCCTAGGCCGGTCGGTCTCGGGCCAC-3'

Nhe-rev 5'-GCGGCTAGCTGCTCGCCCATCGCGGGATGC-3'

The PCR included, in a 50 μl reaction, 5 μl of 10x *Pfu* polymerase buffer (Stratagene), 5 μl 10x z-dNTP mixture (2 mM dATP, 2 mM dCTP, 2 mM dTTP, 1 mM dGTP, 1 mM 7-deaza-GTP), 5 μl DMSO, 2 μl of each primer (10 μM), 1 μl of template DNA (0.1 μg/μl), and 1 μl of cloned *Pfu* polymerase (Stratagene). The PCR conditions were 95°C for 2 min., 25 cycles at 95°C for 30 sec., 60°C for 30 sec., and 72°C for 4 min., followed by 4 min. at 72°C and a hold at 0°C. The amplified DNA products and the Litmus vectors were cut with the appropriate restriction enzymes (*BgI*II and *AvrI*II or *SpeI* and *NheI*), and cloned into either pLitmus 28 or pLitmus38 (New England Bio¹abs), respectively, to generate the constructs designated pKOS60-37-4 and pKOS60-37-2,

30 respectively.

5

10

15

20

25

35

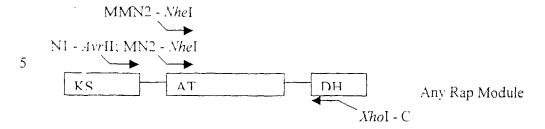
Plasmid pKOS60-29-55 was again used as a template for PCR to amplify sequence 3' to the AT domain using the primers BsrXho-fwd and NsiAfl-rev:

BsrXho-fwd 5'-GATGTACAGCTCGAGTCGGCACGCCCGGCCGCATC-3'
NsiAfl-rev 5'-CGACTCACTTAAGCCATGCATCC-3'

PCR conditions were as described above. The PCR fragment was cut with BsrGI and AfIII, gel isolated, and ligated into pKOS60-37-4 cut with Asp718 and AfIII and

inserted into pKOS60-37-2 cut with *BsrGI* and *AfIII*, to give the plasmids pKOS60-39-1 and pKOS60-39-13, respectively. These two plasmids can be digested with *AvrII* and *XhoI* or *NheI* and *XhoI*, respectively, to insert heterologous AT domains specific for malonyl, methylmalonyl, ethylmalonyl, or other extender units.

- Malonyl and methylmalonyl-specific AT domains were cloned from the rapamycin cluster using PCR amplification with a pair of primers that introduce an *Avr*II or *Nhe*l site at the 5' end and an *Xho*I site at the 3' end. The PCR conditions were as given above and the primer sequences were as follows:
- 10 RATN1 5'-ATCCTAGGCGGGCRGGYGTGTCGTCCTTCGG-3'
 (3' end of Rap KS sequence and universal for malonyl and methylmalonyl CoA),
 RATMN2 5'-ATGCTAGCCGCCGCGTTCCCCGTCTTCGCGCG-3'
 (Rap AT shorter version 5'- sequence and specific for malonyl CoA),
 RATMMN2 5'-ATGCTAGCGGATTCGTCGGTGGTGTTCGCCGA-3'
- (Rap AT shorter version 5'- sequence and specific for methylmalonyl CoA), and RATC 5'-ATCTCGAGCCAGTASCGCTGGTGYTGGAAGG-3' (Rap DH 5'- sequence and universal for malonyl and methylmalonyl CoA).



10

15

Because of the high sequence similarity in each module of the rapamycin cluster, each primer was expected to prime any of the AT domains. PCR products representing ATs specific for malonyl or methylmalonyl extenders were identified by sequencing individual cloned PCR products. Sequencing also confirmed that the chosen clones contained no cloning artifacts. Examples of hybrid modules with the rapamycin AT12 and AT13 domains are shown in a separate figure.

The AvrII-XhoI restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 12 of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below. The AT of rap module 12 is specific for incorporation of malonyl units.

20 AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50 WQLAEALLTLV GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGGCGGC 100 AAVLGHVGGEDIPATAA GTTCAAGGACCTCGGCATCGACTCGCTCACCGCGGTCCAGCTGCGCAACG 150 25 F K D L G I D S L T A V Q L R N CCCTCACCGAGGCGACCGGTGTGCGGCTGAACGCCACGGGGGTCTTCGAC 200 ALTEATGVRLNATAVFD TTCCCGACCCCGCACGTGCTCGCCGGGAAGCTCGGCGACGAACTGACCGG 250 F P T P H V L A G K L G D E L T G 30 CACCCGCGCGCCCGTCGTGCCCCGGACCGCGGCCACGGCCGGTGCGCACG 300 TRAPVVPR TAATAGAH ACGAGCCGCTGGCGATCGTGGGGAATGGCCTGCCGGCTGCCCGGCGGGGGTC 350 EPLAIV G M A C R L P G G V GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400 35 ASPEELWHLVASGTDAI TEFPTDRGWDVDAIYD CGGACCCCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500 PDPDAIGKTEVRHGGFL 40 ACCGGCGCGACAGGCTTCGACGCGCGCGTTCTTCGGCATCAGCCCGCGCGA. 550 TGATGFDAAFFGISPRE GGCCCTCGCGATGGACCCGCAGCAGCGGGTGCTCCTGGAGACGTCGTGGG 600 A L A M D P Q Q R V L L E T S W AGGCGTTCGAAAGCGCCGGCATCACCCCGGACTCGACCCGCGGCAGCGAC 650 45 EAFESAGITPDSTRGSD ACCGGCGTGTTCGTCGGCGCCTTCTCCTACGGTTACGGCACCGGTGCGGA 700 TGVFVGAFSYGYGTGAD CACCGACGGCTTCGGCGCGCCGGCTCGCAGACCAGTGTGCTCTCCGGCC 750 TDGFGATGSOTSVLSG 50 GGCTGTCGTACTTCTACGGTCTGGAGGGTCCGGCGGTCACGGTCGACACG 800 RLSYFYGLEGPAVTVDT GCGTGTTCGTCGTCGCTGGTGGCGCTGCACCAGGCCGGGCAGTCGCTGCG 850 A C S S S L V A L H Q A G Q S L R

	clegaceWilecledeliceaceidaleagageaceicweae.awiaecei	200
	S G E C S L A L V G G V T V M A	
	CTCCCGGCGCTTCGTGGAGTTCTCCCGGCAGCGCGCCTCGCGCCGGAC	950
	S F G G F V E F S R Q R G L A P D	
5	The state of the s	1000
J	GBCCGGGGAAGGCGTTTCGGCGGGGGGGGGACGGCACGAGCTTCGCCGA	1000
	G R A K A F G A G A D G T S F A E	
	GUSTGCCGGTGTGCTGATCGTCGAGAGGCTCTCCGACGCCGAACGCAACG	1050
	G A G V L I V E R L S D A E R N	
		1100
1.0		1100
10	G H T V L A V V R G S A V N Q D G	
	GCCTCCAACGGGCTGTCGGCGCCGAACGGGCCGTCGCAGGAGCGGGTGAT	1150
	ASNGLSAPNGPSQERVI	
	.	. 200
	CCGGCAGGCCTGGCCAACGCCGGGCTCACCCCGGCGGACGTGGACGCCG	1200
	RQALANAGLTPADVDA	
15	TOGAGGOCCACGGCACCGGCACCAGGCTGGGCGACCCCATCGAGGCACAG	1250
	V E A H G T G T R L G D P I E A Q	
		1300
	AVLATYGQERATPLLLG	
	CTCGCTGAAGTCCAACATCGGCCACGCCCAGGCCGCGTCCGGCGTCGCCG	1350
20	S L K S N I G H A Q A A S G V A	
20		1 4 0 0
	GCATCATCAAGATGGTGCAGGCCCTCCGGCACGGGGAGCTGCCGCCGACG	1400
	G I I K M V Q A L R H G E L P P T	
	CTGCACGCGACGAGCGTCGCCGCACGTCGACTGGACGGCCGGC	1450
	L H A D E P S P H V D W T A G A V	
25		1500
23	CGAACTGCTGACGTCGGCCCGGCCGTGGCCCGAGACCGACC	1200
	ELLTSARPWPETDRPR	
	GGGCAGGCGTGTCGTCCTTCGGGATCAGTGGCACCAACGCCCACGTCATC	1550
	R A G V S S F G I S G T N A H V I	
		1.000
3.0	CTGGAAAGCGCACCCCCACTCAGCCTGCGGACAACGCGGTGATCGAGCG	7000
30	LESAPPTQPADNAVIER	
	GGCACCGGAGTGGGTGCCGTTGGTGATTTCGGCCAGGACCCAGTCGGCTT	1650
	A P E W V P L V I S A R T Q S A	
		1700
		2.00
2.5	L T E H E G R L R A Y L A A S P G	
35	GTGGATATGCGGGCTGTGGCATCGACGCTGGCGATGACACGGTCGGT	1750
	V D M R A V A S T L A M T R S V F	
	CGAGCACCGTGCCGTGCTGGGAGATGACACCGTCACCGGCACCGCTG	1800
	E H R A V L L G D D T V T G T A	
		1050
	TGTCTGACCCTCGGGCGGTGTTCGTCTTCCCGGGACAGGGGTCGCAGCGT	-820
40	V S D P R A V F V F P G Q G S Q R	
	GCTGGCATGGGTGAGGAACTGGCCGCCGCGTTCCCCCGTCTTCGCGCGGAT	1900
	A G M G E E L A A A F P V F A R I	
	A G M G E E E A A A F F V F A A F F V F A A F F V F A A F F V F A A F F A A F F F A A F F A A F F F A A F F A A F F A A F F F A A F F A A F F A A F F A A F F A A F F A A F F A A F F A A F F A A F F A A F A F A A F A F A A F A F A A F A F A F A A F A A F A F A A F A F A A F A A F A A F A A F A A F A A F A A F A A F	1050
	CCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCGATCTGGAGGTGAACG	1950
	HQQVWDLLDVPDLEVN	
45	AGACCGGTTACGCCCAGCCGGCCCTGTTCGCAATGCAGGTGGCTCTGTTC	2000
	E T G Y A Q P A L F A M Q V A L F	
		2050
	GGGCTGCTGGAATCGTGGGGTGTACGACCGGACGCGGTGATCGGCCATTC	2050
	G L L E S W G V R P D A V I G H S	
	GGTGGGTGAGCTTGCGGCTATGTGTCCGGGGTGTGGTCGTTGGAGG	2100
50	V G E L A A A Y V S G V W S L E	
50		0.550
	ATGCCTGCACTTTGGTGTCGGCGCGGGCTCGTCTGATGCAGGCTCTGCCC	2150
	D A C T L V S A R A R L M Q A L P	
	GCGGGTGGGGTGATGGTCGCTGTCCCGGTCTCGGAGGATGAGGCCCGGGC	2200
	A G G V M V A V P V S E D E A R A	
<i></i>		0050
55	CGTGCTGGGTGAGGTGTGGAGATCGCCGCGGTCAACGGCCCGTCGTCGG	2250
	V L G E G V E I A A V N G P S S	
	TGGTTCTCTCCGGTGATGAGGCCGCCGTGCTGCAGGCCGCGGAGGGGCTG	2300
	V V L S G D E A A V L Q A A E G L	
		0256
	GGGAAGTGGACGCGCTGGCGACCAGCCACGCGTTCCATTCCGCCCGTAT	2350
60	G K W T R L A T S H A F H S A R M	
	GGAACCCATGCTGGAGGAGTTCCGGGCGGTCGCCGAAGGCCTGACCTACC	2400
	E P M L E E F R A V A E G L T Y	
		2450
	GGACGCCGCAGGTCTCCATGGCCGTTGGTGATCAGGTGACCACCGCTGAG	2450
	RTPQVSMAVGDQVTTAE	

	TACTGGGTGCGGCAGGTCCGGGACACGGTCCGGTTCGGCCGAGCAGGTGGC	2500
	Y,W V R Q V R D T V R F G E Q V A	
	CTOGTACGAGGACGCCGTCTTCGTCGAGCTGGGTGCCGACCGGTCACTGG	2550
5	S Y E D A V F V E L G A D R S L DOGGCTGGTCGACGGTGTCGCGACGACGACGAAATCCAG	2600
5	A R L V D G V A M L H G D H E I Q	2000
	GCCGCGATCGGCGCCCTGGCCCACCTGTATGTCAACGGCGTCACGGTCGA	2650
	A A I G A L A H L Y V N G V T V D	
	CTGGCCGCGCTCCTGGGCGATGCTCCGGCAACACGGGTGCTGGACCTTC	2700
10	W P A L L G D A P A T R V L D L	
	CGACATACGCCTTCCAGCACCAGCGCTACTGSCTCGAGTCGGCACGCCCG	2750
	PTYAFQHQRYWLESARP	0000
		2800
15	A A S D A G H P V L G S G I A L A cogsatoscomponences of the costscomposition	2850
	G U . R V F T G S V P T G S V	2.000
	GCGCCCTGTTCGTCGCCGAGCTCGCCGCTGGCCGCCGCGGACGCGGTCGAC	2900
	RAVFVAELALAAADAVD	
	TGCGCCACGGTCGAGCGGCTCGACATCGCCTCCGTGCCCGGCCGG	2950
20	C A T V E R L D I A S V P G R P G	5000
	CCATGGCCGGACGACCGTACAGACCTGGGTCGACGAGCCGGCGGACGACG	3000
	H G R T T V Q T W V D E P A D D GCCGGCGCGGGGGGCGGTGGACG	3050
	G R R R F T V H T R T G D A P W T	2020
25	OTGCACGCGAGGGGGTGCTGCGCCCCATGGCACGGCCCTGCCCGATGC	3100
	L H A E G V L R P H G T A L P D A	
	GGCCGACGCCGAGTGGCCCCCACCGGGCGCGGTGCCCGCGGACGGGCTGC	3150
	A D A E W P P P G A V P A D G L	
20	CGGGTGTGTGGCGCGGGGGGGCCAGGTCTTCGCCGAGGCCGAGGTGGAC	3200
30	P G V W R R G D Q V F A E A E V D	2250
	GGACCGGACGGTTTCGTGGTGCACCCCGACCTGCTCGACGCGGTCTTCTC G P D G F V V H P D L L D A V F S	3250
	G P D G F V V H P D L L D A V F S CGCGGTCGGCGACGGAGCCGCCGGCCGGATGGCGGACCTGACGG	3300
	A V G D G S R Q P A G W R D L T	5500
35	TGCACGCGTCGGACGCCACCGTACTGCGCGCCTGCCTCACCCGGCGCACC	3350
	V H A S D A T V L R A C L T R R T	
	GACGGAGCCATGGGATTCGCCGCCTTCGACGGCGCCGGCCTGCCGGTACT	3400
	D G A M G F A A F D G A G L P V L	2150
40	CACCGCGGAGGCGGTGACGCTGCGGGAGGTGGCGTCACCGTCCGGCTCCG	3450
40	T A E A V T L R E V A S P S G S AGGAGTCGGACGGCTGCACCGGTTGGAGTGGCTCGCGGTCGCCGAGGCG	3500
	E E S D G L H R L E W L A V A E A	
	GTCTACGACGGTGACCTGCCCGAGGGACATGTCCTGATCACCGCCGCCCA	3550
	V Y D G D L P E G H V L I T A A H	
45	CCCCGACGACCCCGAGGACATACCCACCCGCGCCCACACCCGCGCCCACCC	3600
	PDDPEDIPTRAHTRAT	
	GCGTCCTGACCGCCCTGCAACACCACCTCACCACCACCGACCACCCTC	3650
	R V L T A L Q H H L T T 1 D H T L	3700
50	ATCGTCCACCACCACCGCCGCCGCCGCCCCCGTCACCGGCCTCAC I V H T T T D P A G A T V T G L T	3700
50	CCGCACCGCCCAGAACGAACACCCCCACCGCATCGGCCTCATCGAAACCG	3750
	R T A Q N E H P H R I R L I E T	
	ACCACCCCACACCCCCTCCCCCTGGCCCAACTCGCCACCCTCGACCAC	3800
	DHPHTPLPLAQLATLDH	
55	CCCCACCTCCGCCTCACCCACCACCACCCCCCCCCCCCC	3850
	PHLRLTHHTLHHPHLTP	
	CCTCCACACCACCACCACCACCACCACCCCCTCAACCCCGAACACG	3900
	L H T T T P P T T T P L N P E H	3050
60	CCATCATCACCGGCGGCTCCGGCACCCTCGCCGGCATCCTCGCCCGC A I I I T G G S G T L A G I L A R	2320
UU	A I I I T G G S G T L A G I L A R CACCTGAACCACCCCCACACCTACCTCTCTCCCGCACCCCACCCCCCGA	4000
	H L N H P H T Y L L S R T P P P D	.000
	CGCCACCCCGGCACCCACCTCCCTGCGACGTCGGCGACCCCCCCC	4050
	A T P C T U T P C D V G D P H O	

TCGCCACCACCCTCACCCACATCCCCCAACCCCTCACGGCCATCTTCCAC 4100 LATTLTHIPQPLTAIFH ACCGCCGCCACCCTCGACGACGGCATCCTCCACGCCCTCACCCCCGACCG 4150 TAATLDDGILHALTPDR 5 CCTCACCACGTCCTCCACCGCARAGCCAACGCCGCCTGGCACCTGCACC 4200 LTTVLHPKANAAWELE ACCTCACCCAAAACCAACCCCTCACCCACTTCGTCCTCTACTCCAGCGCC 4250 H L T Q N Q P L T H F V L Y S S A GCCGCCGTCCTCGGCAGCCCGGACAAGGAAACTACGCCGCCGCCAACGC 4300 10 A A V L G S P G Q G N Y A A A N A CTTCCTCGACGCCTCGCCACCCCACCGCCACACCCTCGGCCAACCCGCCA 4350 F L D A L A T H R H T L G Q P A CCTCCATCGCCTGGGGCATGTGGCACACCACCAGCACCCTCACCGGACAA 4400 T S I A W G M W H T T S T L T G Q 15 CTCGACGACGCCGACCGGGACCGCATCCGCCGCGGGGGGTTTCCTCCCGGAT 4450 LDDADRDRIRRGGFLPI CACGGACGACGAGGGCATGGGATGCAT T D D E G

The AvrII-XhoI restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 13 (specific for methylmalonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below.

AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50 25 Q L A E A L L T L V R E S T GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGGCGGC 100 A A V L G H V G G E D I P A T A A GTTCAAGGACCTCGGCATCGACTCGCTCACCGCGGTCCAGCTGCGCAACG 150 F K D L G I D S L T A V Q L R N 30 CCCTCACCGAGGCGACCGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200 A L T E A T G V R L N A T A V F D TTCCCGACCCGCACGTGCTCGCCGGGAAGCTCGGCGACGAACTGACCGG 250 F P T P H V L A G K L G D E L T G CACCGGGGGCGCGTCGTGCCCGGGACCGCGGGCCACGGCCGGTGCGCACG 300 35 TRAPVVPRTAATAGAH ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGGCTGCCCGGCGGGGTC 350 DEPLAIVGMACRLPGGV GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400 ASPEELWHLVASGTDAI 40 CACGGAGTTCCCGACGGACCGCGGGGGGGGACGTCGACGCGATCTACGACC 450 T E F P T D R G W D V D A I Y D CGGACCUCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500 P D P D A I G K T F V R H G G F L ACCGGCGCGACAGGCTTCGACGCGGCGTTCTTCGGCATCAGCCCGCGCGA 550 45 T G A T G F D A A F F G T S P R E GGCCCTCGCGATGGACCCGCAGCAGCGGGTGCTCCTGGAGACGTCGTGGG 600 A L A M D P Q Q R V L L E T S W AGGCGTTCGAAAGCGCCGGCATCACCCCGGACTCGACCCGCGGCAGCGAC 650 E A F E S A G I T P D S T R G S D 50 ACCGGCGTGTTCGTCGGCGCCTTCTCCTACGGTTACGGCACCGGTGCGGA 700 G V F V G A F S Y G Y G T G A D CACCGACGGCTTCGGCGCGACCGGCTCGCAGACCAGTGTGCTCTCCGGCC 750 T D G F G A T G S O T S V L S G GGCTGTCGTACTTCTACGGTCTGGAGGGTCCGGCGGTCACGGTCGACACG 800 55 R L S Y F Y G L E G P A V GCGTGTTCGTCGTCGCTGGTGGCGCTGCACCAGGCCGGGCAGTCGCTGCG 850 A C S S S L V A L H Q A G Q S L R CTCCGGCGAATGCTCGCTCGCCCTGGTCGGCGGCGTCACGGTGATGGCGT 900 LALVGGVTVMA SGECS CTCCCGGCGGCTTCGTGGAGTTCTCCCGGCAGCGCGGCCTCGCGCCGGAC 950 60

SPGGFVEFSRQRGLAPD GGCCGGGCGAAGGCGTTCGGCGCGGGGGGGGGGGGGGGACGAGCTTCGCCGA 1000 G R A K A F G A G A D G T S F A E GGGTGCCGGTGTGCTGATCGTGSAGAGGCTCTCCGACGCCGAACGCAACG 1050 G A G V L I V E R L S D A E R N GTOACACCGTCCTGGCGGTCGTCCGTCGGCGGGGGTCAAACCAGGATCGT 1100 G H T V L A V V R G S A V N O D G GCCTCCHACGGGCTGTCGGCGCCGAACGGGCCGTCGCAGGAGGGGGTGAT 1150 A S N G L S A P N G P S Q E R V I 10 CCGGCAGGCCCTGGCCAACGCCGGGCTCACCCCGGCGGACGTGGACGCCG 1200 R Q A L A N A G L T P A D V D A TCGAGGCCCACGGCACCGGCACCAGGCTGGGCGACCCCATCGAGGCACAG 1250 V E A H G T G T R L G D P I E A Q GCGGTACTGGCCACCTACGGACAGGASCGCGCCACCCCCCCTGCTGCTGGG 1300 15 AVLATYGOERATPLLLG OTOGOTGAAGTOCAAGATQGQCCAGGCCAAGGCGGGGTCGGGGTCGGCGG 1350 SLKSNIGHAQAASGVA GCATCATCAAGATGGTGCAGGCCCTCCGGCACGGGGAGSTGCCGCCGACG 1400 GIIKMVQALRHGELPPT 20 LHADEPSPHVDWTAGAV ELLTSARPWPETDRPR GGGCGGGCGTGTCGTCCTTCGGAGTCAGCGGCACCAACGCCCACGTCATC 1550 25 R A G V S S F G V S G T N A H V I CTGGAGAGCGCACCCCCGCTCAGCCCGCGGAGGAGGCGCAGCCTGTTGA 1600 L E S A P P A Q P A E E A Q P V E GACGCCGGTGGTGGCCTCGGATGTGCTGCCGCTGGTGATATCGGCCAAGA 1650 T P V V A S D V L P L V I S A K 30 CCCAGCCCGCCTGACCGAACACGAAGACCGGCTGCGCGCCTACCTGGCG 1700 TOPALTEHEDRLRAYLA GCGTCGCCCGGGGCGGATATACGGGCTGTGGCATCGACGCTGGCGGTGAC 1750 A S P G A D I R A V A S T L A V T ACGSTCGGTGTTCGAGCACCGCGCCGTACTCCTTGGAGATGACACCGTCA 1800 35 RSVFEHRAVLLGDDTV CCGGCACCGGGTGACCGACCCCAGGATCGTGTTTGTCTTTCCCGGGCAG 1850 TGTAVTDPRIVFVFPGQ GGGTGGCAGTGGCTGGGGATGGGCAGTGCACTGCGCGATTCGTCGGTGGT 1900 G W Q W L G M G S A L R D S S V V 40 GTTCGCCGAGCGGATGGCCGAGTGTGCGGCGGCGTTGCGCGAGTTCGTGG 1950 F A E R M A E C A A A L R E F V ACTGGGATCTGTTCACGGTTCTGGATGATCCGGCGGTGGTGGACCGGGTT 2000 D W D L F T V L D D P A V V D R V GATGTGGTCCAGCCCGCTTCCTGGGCGATGATGGTTTCCCTGGCCGCGGT 2050 45 D V V Q P A S W A M M V S L A A V GTGGCAGGCGGCCGGTGTGCGGCCGGATGCGGTGATCGGCCATTCGCAGG 2100 W Q A A G V R P D A V I G H S Q GTGAGATCGCCGCAGCTTGTGTGGCGGGTGCGGTGTCACTACGCGATGCC 2150 G E I A A A C V A G A V S L R D A 50 GCCCGGATCGTGACCTTGCGCAGCCAGGCGATCGCCCGGGGCCTGGCGGG 2200 ARIVTLRSQAIARGLAG CCGGGGCGCGATGGCATCCGTCGCCCTGCCCGCGCAGGATGTCGAGCTGG 2250 R G A M A S V A L P A Q D V E L TCGACGGGGCCTGGATCGCCGCCCACAACGGGCCCGCCTCCACCGTGATC 2300 55 V D G A W I A A H N G P A S T V I GCGGGCACCCGGAAGCGGTCGACCATGTCCTCACCGCTCATGAGGCACA 2350 A G T P E A V D H V L T A H E A Q AGGGGTGCGGGTGCGGCGATCACCGTCGACTATGCCTCGCACACCCCGC 2400 60 ACGTCGAGCTGATCCGCGAGGAGCTACTCGACATCACTAGCGACAGCAGC 2450 H V E L I R D E L L D I T S D S S TCGCAGACCCCGCTCGTGCCGTGGCTGTCGACCGTGGACGGCACCTGGGT 2500 SQTPLVPWLSTVDGTWV CGACAGCCGCTGGACGGGGAGTACTGGTACCGGAACCTGCGTGAACCGG 2550

D S P L D G E Y W Y R N L R E P TCGGTTTCCACCCGCCGTCAGCCAGTTGCAGGCCCAGGGCGACACCGTG 2600 V G F H P A V S Q L Q A Q G D T V TTCGTCGAGGTCAGCGCCAGCCGGGTGTTGTTGCAGGCGATGGACGACGA 2650 F V E V S A S P V L L Q A M D D D TGTCGTCACGGTTGCCACGCTGCGTCGTGACGACGGCGACGCCACCCGGA 2700 V V T V A T L R R E D G D A T R TGCTCACCCCCTGCCACAGGCCTATGTCCACGGCGTCACCGTCGACTGG 2750 M L T A L A Q A Y V H G V T V D W 10 CCCGCCATCCTCGGCACCACCACACCCGGGTACTGGACCTTCCGACCTA 2800 PAILGTTTTRVLDLPTY CGCCTTCCAACACCAGCGGTACTGGCTCGAGTCGGCACGCCCGGCCGCAT 2850 A F Q H Q R Y W L E S A R P A A CCGACGCGGGCCACCCCGTGCTGGGCTCCGGTATCGCCCTCGCCGGGTCG 2900 15 S D A G H P V L G S G I A L A G S CCGGGGGGGGTGTTCACGGGTTCCGTGCGACCGGTGGGGACCGCGGGT 2950 PGRVFTGSVPTGADRAV F V A E L A L A A A D A V D C A 20 T V E R L D I A S V P G R P G H G R T T V Q T W V D E P A D D G R R CCGGTTCACCGTGCACACCCGCACCGGCGCGCGCGTGGACGCTGCACG 3150 25 R F T V H T R T G D A P W T L H CCGAGGGGGTGCTGCCCCCATGGCACGGCCCTGCCCGATGCGGCCGAC 3200 A E G V L R P H G T A L P D A A D GCCGAGTGGCCCCCACCGGGCGCGGTGCCCGCGGACGGGCTGCCGGGTGT 3250 A E W P P P G A V P A D G L P G V 30 W R R G D Q V F A E A E V D G P ACGGTTTCGTGGTGCACCCCGACCTGCTCGACGCGGTCTTCTCCGCGGTC 3350 D G F V V H P D L L D A V F S A V GGCGACGGAAGCCGCCAGCCGGCCGGATGGCGCGACCTGACGGTGCACGC 3400 35 G D G S R Q P A G W R D L T V H A GTCGGACGCCACCGTACTGCGCGCCTGCCTCACCCGGCGCACCGACGGAG 3450 S D A T V L R A C L T R R T D G CCATGGGATTCGCCGCCTTCGACGGCGCCGGCCTGCCGGTACTCACCGCG 3500 AMGFAAFDGAGLPVLTA 40 GAGGCGGTGACGCTGCGGGAGGTGGCGTCACCGTCCGGCTCCGAGGAGTC 3550 EAVTLREVASPSGSEES GGACGGCCTGCACCGGTTGGAGTGGCTCGCGGTCGCCGAGGCGGTCTACG 3600 D G L H R L E W L A V A E A V Y ACGGTGACCTGCCCGAGGGACATGTCCTGATCACCGCCGCCCCACCCCGAC 3650 D G D L P E G H V L I T A A H P D 45 GACCCCGAGGACATACCCACCCGCGCCCACACCCGCGCCACCCGCGTCCT 3700 D P E D I P T R A H T R A T R V L GACCGCCCTGCAACACCACCACCACCACCACCACCACCCTCATCGTCC 3750 TALQHHLTTTDHTLIV 50 ACACCACCACCGACCCGCCGGCGCCCACCGTCACCGGCCTCACCCGCACC 3800 H T T T D P A G A T V T G L T R T GCCCAGAACGAACACCCCCACCGCATCCGCCTCATCGAAACCGACCACCC 3850 AQNEHPHRIRLIETDHP CCACACCCCCTCCCCTGGCCCAACTCGCCACCCTCGACCACCCCCACC 3900 55 HTPLPLAQLATLDHPH TCCGCCTCACCCACCACCCTCCACCACCCCCACCTCACCCCCCTCCAC 3950 LRLTHHTLHHPHLTPLH ACCACCACCCACCACCACCACCACCCCCTCAACCCCGAACACGCCATCAT 4000 TTTPPTTTPLNPEHAII 60 ITGGSGTLAGILARHL ACCACCCCACACCTACCTCTCTCCCGCACCCCACCCCCGACGCCACC 4100 NHPHTYLLSRTPPPDAT CCCGGCACCCACCTCCCCTGCGACGTCGGCGACCCCCACCAACTCGCCAC 4150

P G T H L P C D V G D P H Q L A CACCCTCACCCACATOCCCCAACCCCTCACCGCCCATCTTCCACACCCCCG 4200 TLL T H I P Q P L T A I F H T A CCACCCTCGACGACGCATCCTCCACGCCCTCACCCCGACCGCCTCACC 4250 ATLDDGILHALTPDRLT TVLHPKANAAWHLHHLT CCAAAACCAACCCCTCACCCACTTCGTCCTCTACTCCAGCGCCGCCGCCGCCG 4350 O N O P L T H F V L Y S S A A A 10 TOOTOGGCAGOCCCGGACAAGGAAACTACGCCGCCCAACGCCTTCCTC 4400 V L G S P G Q G N Y A A A N A F L GACGCCTCGCCACCCACCCCCACACCCTCGGCCAACCCGCCACCTCCAT 4450 DALATHRHTLGQPATSI CGCCTGGGGCATGTGGCACACCACCAGCACCTCACCGGACAACTCGACG 4500 15 A W G M W H T T S T L T G Q L D ACCCCCACCGGGGACCGCATCCGCCGCGGGCGGTTTCCTCCCGATCACGGAC 4550 DADRDRIRR GFLFITD GACGAGGO L'GGGGATGCAT D E G 20

The *Nhe*II-*Xho*I restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 12 (specific for malonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below.

25 AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50 Q L A E A L L T L V R E S T GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGCGGC 100 A A V L G H V G G E D I P A T A A GTTCAAGGACCTCGCCATCGACTCGCTCACCGCGGTCCAGCTGCGCAACG 150 30 F K D L G I D S L T A V Q L R N CCCTCACCGAGGCGACCGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200 ALTEATGVRLNATAVFD TTCCCGACCCGCACGTGCTCGCCGGGGAAGTTCGGCGACGAACTGACCGG 250 F P T P H V L A G K L G D E L T G 35 CACCCGCGCGCCCGTCGTGCCCCGGACCGCCGCCACGGCCGGTGCGCACG 300 TRAPVVPRTAATAGAH ACGAGCCGCTGGCGATCGTGGGGAATGGCCTGCCGGCTGCCCGGCGGGGGTC 350 DEPLAIVGMACRLPGGV GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400 40 ASPEELWHLVASGTDAI TEFPTDRGWDVDAIYD CGGACCCCGACGCGATCGGCAGGACCTTCGTCCGGCACGGTGGCTTCCTC 500 PDPDAIGKTFVRHGGFL 45 ACCGGCGCGACAGGCTTCGACGCGGCGTTCTTCGGCATCAGCCCGCGCGA 550 TGATGFDAAFFGISPRE GGCCCTCGCGATGGACCCGCAGCAGCGGGTGCTCCTGGAGACGTCGTGGG 600 ALAMDPOORVLLETSW AGGCGTTCGAAAGCGCCGGCATCACCCCGGACTCGACCCGCGCCAGCGAC 650 50 E A F E S A G I T P D S T R G S D ACCGGCGTGTTCGTCGGCGCCTTCTCCTACGGTTACGGCACCGGTGCGGA 700 TGVFVGAFSYGYGTGAD CACCGACGGCTTCGGCGCGACCGGCTCGCAGACCAGTGTGCTCTCCGGCC 750 T D G F G A T G S Q T S V L S G 55 GGCTGTCGTACTTCTACGGTCTGGAGGGTCCGGCGGTCACGGTCGACACG 800 RLSYFYGLEGPAVTVDT GCGTGTTCGTCGTCGCTGGTGGCGCTGCACCAGGCCGGGCAGTCGCTGCG 850 ACSSSLVALHQAGQSLR CTCCGGCGAATGCTCGCTCGCCCTGGTCGGCGGCGTCACGGTGATGGCGT 900 60 SGECSLALVGGVTVMA

CTCCCGGCGGCTTCGTGGAGTTCTCCCGGCAGCGGCGCCTCGCGCCGGAC 950 S P G G F V E F S R Q R G L A P D G R A K A F G A G A D G T S F A E GSGTSCCGGTGTGCTGATCGTCGAGAGSCTCTCCGACGCCGAACGCAACG 1050 G A G V L I V E R L S D A E R N GTCACACCGTCCTGGCGGTCGTCCGTGGTTCGGCGGTCAACCAGGATGGT 1100 G H T V L A V V R G S A V N Q D G GCCTCUAACGCGCTCTCGGCGCCGAACGGGCCGTCGCAGGAGCGGGTGAT 1150 10 ASNGLSAPNGPSQERVI CCGGCAGGCCCTGGCCAACGCCGGGGTCACCCCGGCGGACGTGGACGCCG 1200 R Q A L A N A G L T P A D V D A TCGAGGCCCACGGCACCGGCACCAGGCTGGGCGACCCCATCGAGGCACAG 1250 V E A H G T G T R L G D P I E A Q 15 SOGGTACTGGCCACCTACGGACAGGAGCGCGCCACCCCCTGCTGCTGGG 1300 A V L A T Y G Q E R A T P L L L G OTOGOTALA E E ELAJAMOGGCOACGCCAGGCCGCGTCGGCGCGTCGCAG S 1 K s N I G H A Q A A S G v A 20 G I I K M V Q A L R H G E L P P T L H A D E P S P H V D W T A G A V ELLTSARPWPETDRPR 25 GTGCCGCCGTCTCCTCGTTCGGGGTGAGCGCCACCAACGCCCACGTCATC 1550 R A A V S S F G V S G T N A H V I CTGGAGGCCGGACCGGTAACGGAGACGCCCGCGGGCATCGCCTTCCGGTGA 1600 L E A C P V T E T P A A S P S G D CCTTCCCCTGCTGGTGTCGGCACGCTCACCGGAAGCGCTCGACGAGCAGA 1650 30 LPLLVSARSPEALDEQ TCCGCCGACTGCGCCTACCTGGACACCCCCGGACGTCGACCGGGTG 1700 I R R L R A Y L D T T P D V D R V GCCGTGGCACAGACGCTGGCCCGGCGCACACTTCGCCCACCGCGCCGT 1750 AVAOTLARRTHFAHRAV 35 GCTGCTCGGTGACACCGTCATCACCACACCCCCGGGGACCGGCCCGACG 1800 LLGDTVITTPPADRPD AACTOGTOTTCGTCTACTCCGGCCAGGGCACCCAGCATCCCGCGATGGGC 1850 E L V F V Y S G Q G T Q H P A M G GAGCAGCTAGCCGCCGCGTTCCCCGTCTTCGCGCGGGATCCATCAGCAGGT 1900 40 EQLAAAFPVFARIHQQV GTGGGACCTGCTCGATGTGCCCGATCTGGAGGTGAACGAGACCGGTTACG 1950 WDLLDVPDLEVNETGY CCCAGCCGGCCCTGTTCGCAATGCAGGTGGCTCTGTTCGGGCTGCTGGAA 2000 A Q P A L F A M Q V A L F G L L E 45 S W G V R P D A V I G H S V G E L TGCGGCTGCGTATGTCCGGGGTGTGGTCGTTGGAGGATGCCTGCACTT 2100 A A A Y V S G V W S L E D A C T TGGTGTCGGCGCGGGCTCTGTCTGATGCAGGCTUTGCCCLUGGGTGGGGTG 2150 50 L V S A R A R L M Q A L P A G G V ATGGTCGCTGTCCCGGTCTCGGAGGATGAGGCCCGGGCCGTGCTGGGTGA 2200 M V A V P V S E D E A R A V L G E GGGTGTGGAGATCGCCGCGGTCAACGGCCCGTCGTCGGTGGTTCTCTCCG 2250 G V E I A A V N G P S S V V L S 55 GTGATGAGGCCGCGTGCTGCAGGCCGCGGAGGGGCTGGGGAAGTGGACG 2300 G D E A A V L Q A A E G L G K W T CGGCTGGCGACCACGCGTTCCATTCCGCCCGTATGGAACCCATGCT 2350 R L A T S H A F H S A R M E P M L GGAGGAGTTCCGGGCGGTCGCCGAAGGCCTGACCTACCGGACGCCGCAGG 2400 60 E E F R A V A E G L T Y R T P Q TCTCCATGGCCGTTGGTGATCAGGTGACCACCGCTGAGTACTGGGTGCGG 2450 V S M A V G D Q V T T A E Y W V R CAGGTCCGGGACACGGTCCGGTTCGGCGAGCAGGTGGCCTCGTACGAGGA 2500 Q V R D T V R F G E Q V A S Y E D

	CGCCGTGTTCGTCGAGCTGGGTGCCGACCGGTCACTGGCCCGCCTGGTCG	2550
	A V F V E L G A D R S L A R L V	2600
	ACGSTGTCGCGATGCTGCACGGCGACGACGAGAATCCAGGCCGCGATCGGC D G V A M L H G D H E I Q A A I G	2000
5	GOODDOODDOODDOODDOODDOODDOODDOODDOODDOO	2650
	A L A H L Y V N G V T V D W P A L	
	COTGGGCGATGCTCCGGCAACACGGGTGCTGGACCTTCCGACATACGCCT	2700
	L G D A P A T R V L D L P T Y A	0750
10	TCCAGCAGCAGCGCTACTGGCTCGAGTCGGCACGCCCGGCCGCATCCGAC F O H O R Y W L E S A R P A A S D	2750
10		2800
	A G H P V L G S G I A L A G S P G	2000
		2850
	RVFTGSVPTGADRAVF	
15	TOGCOGAGOTGGCCGCCGCCGGGACGGGTCGACTGCGCCACGGTC	2900
	V A E L A L A A A D A V D C A T V	2950
	E R L D I A S 1 P G R P G H G R T	2230
	GACCGTACAGACCTGGGTCGACGACGACGACGACGACGGCGGCGGCGGCGGCGGCGGC	3000
20	TVQTWVDEPADDGRRR	
	TCACCGTGCACACCCGCACCGGCGACGCCCGTGGACGCTGCACGCCGAG	3050
	F T V H T R T G D A P W T L H A E	3100
	GGGGTGCTGCCCCATGGCACGCCCGATGCGGCCGACGCCCGACGCCGACGCCCGACGCCCGACGCCCGACGCCCGACGCCCGACGCCCGACGCCCGACGCCCGACGCCCGACGCCCGACGCCCGACGCCCGACGCCCGACGCCCGACGCCCGACGCCCGACGCCCGACGCCCGACGCCACACGCCCACACACACACACACACACACACACACACACACACACA	2100
25	GTGGCCCCACCGGGCGCGGGTGCCCGCGGACGGGCTGCCGGGTGTGTGGC	3150
	W P P P G A V P A D G L P G V W	
	GCCGGGGGACCAGGTCTTCGCCGAGGCCGAGGTGGACGGAC	3200
	R R G D Q V F A E A E V D G P D G TTCGTGGTGCACCCGACCTGCTCGACGGGTCTTCTCCGCGGTCGGCGA	3250
30	F V V H P D L I D A V F S A V G D	3230
	CGGAAGCCGCCAGCCGGCCGGATGGCGCGACCTGACGGTGCACGCGTCGG	3300
	G S R Q P A G W R D L T V H A S	
	ACGCCACCGTACTGCGCGCCCTGCCTCACCCGGCGCACCGACGGAGCCATG	3350
35	D A T V L R A C L T R R T D G A M GGATTCGCCGCCTTCGACGGCGCCGGCCTGCCGGTACTCACCGCGGAGGC	3400
33	G F A A F D G A G L P V L T A E A	0.00
	GGTGACGCTGCGGGAGGTGGCGTCACCGTCCGGCTCCGAGGAGTCGGACG	3450
	V T L R E V A S P S G S E E S D	2500
40		3500
40	G LOH R L E W L A V A E A V Y D G GACCTGCCCGAGGGACATGTCCTGATCACGGCGCCCCACCCCGACGACGACCCC	3550
	D L P E G H V L I T A A H P D D P	
	CGAGGACATACCCACCCGCGCCCACACCCGCGCCACCCGCGTCCTGACCG	3600
15	E D I P T R A H T R A T R V L T	2650
45	CCCTGCAACACCACCTCACCACCACCACCACCACCCTCATCGTCCACACC	2020
	ACCACCGACCCGCCGCCGCCACCGTCACCGGCCTCACCGCCACCGCCA	3700
	T T D P A G A T V T G L T R T A Q	
50	GAACGAACACCCCCACCGCATCCGCCTCATCGAAACCGACCACCCCCCACA	3750
50	N E H P H R I R L I E T D H P H CCCCCCTCCCCCTGGCCCAACTCGCCACCCTCGACCACCCCCCACCTCGCC	3800
	T P L P L A Q L A T L D H P H L R	5000
	CTCACCACCACCACCACCACCCCACCTCACCCCCCCCCC	3850
	LTHHTLHHPHLTPLHTT	
55	CACCCCACCCACCACCACCCCCCTCAACCCCGAACACGCCATCATCA	3900
	T P P T T T P L N P E H A I I I COGGOGGOTOCGGCACCTCGCCGGCATCCTCGCCGCCACCTGAACCAC	3950
	T G G S G T L A G I L A R H L N H	3730
	CCCCACACCTACCTCCTCTCCCGCACCCCACCCCCGACGCCACCCCCGG	4000
60	PHTYLLSRTPPPDATPG	
	CACCCACCTCCCCTGCGACGTCGGCGACCCCCACCAACTCGCCACCACCC	4050
	CACCCACCTCCCCTGCGACGTCGGCGACCCCACCACCTCGCCACCACCCC T H L P C D V G D P H Q L A T T TCACCCACATCCCCCAACCCCTCACCGCCATCTTCCACACCGCCGCCACC	

CTCGACGACGCCATCCTCCACGCCCTCACCCCCCGACCGCCTCACCACCGT 4150 LDDGILHALTPDRLTTV CCTCCACCCCAAAGCCAACGCCGCCTGGCACCTGCACCACCTCACCCAAA 4200 LHPKANAAWHLHHLTQ ACCRACCOCTOACCOACTTCGTCCTCTACTCCAGCGCCGCCGCCGTCCTC 4250 N Q P L T H F V L Y S S A A A V L GGCAGCCCGGACAAGGAAACTACGCCGCCGCCAACGCCTTCCTCGACGC 4300 G S P G Q G N Y A A A N A F L D A CCTGGCCACGCCACGCCACACCCTGGGCCAAGCGGCCACGTCGATGGCCT 4350 10 LATHRHTLGOPATSIA GGGGCATGTGGCACACCACCACCACCCTCACCGGACAACTCGACGACGCC 4400 W G N W H T T S T L T G Q L D D A GACCGGGACCGCATCCGCCGCGGGGGGTTTCCTCCCGATCACGGACGACGACA 4450 D R D R I R R G G F L P I T D D E 15 GGGCATGGGGATGCAT

20

The *NheII-XhoI* restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 13 (specific for methylmalonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below.

AGATOTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50 QLAEALLTLVREST GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGGCGGC 100 25 A A V L G H V G G E D I P A T A A GTTCAAGGACCTCGGCATCGACTCGCTCACCGCGGTCCAGCTGCGCAACG 150 F K D L G I D S L T A V Q L R N CCCTCACCGAGGGGACCGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200 ALTEATGVRLNATAVED 30 TTCCCGACCCGCACGTGCTCGCCGGGAAGCTCGGCGACGAACTGACCGG 250 F P T P H V L A G K L G D E L T G CACCCGCGCGCCCGTCGTGCCCCGGACCGCGGCCACGGCCGGTGCGCACG 300 TRAPVVPRTAATAGAH ACGASCOGOTGGCGATCGTGGGGAATGGCCTGCCGGCTGCCCGGCGGGGTC 350 35 DEPLAIV G M A C R L P G G V GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400 ASPEELWHLVASGTDA CACGGAGTTCCCGACGGACGCGGCTGGGACGTCGACGCGATCTACGACC 450 TEFPTDRGWDVDAIYD 40 CGGACCCCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500 DPDAIGKTFVRHGGFL ACCGGCGCGACAGGCTTCGACGCGGCGTTCTTCGGCATCAGCCCGCGCGA 550 TGATGFDAAFFGISPRE GGCCCTCGCGATGGACCCGCAGCAGCGGGTGCTCCTGGAGACGTCGTGGG 600 45 A L A M D P Q Q R V L L E T S W AGGCSTTCGAAAGCGCCGGCATCACCCCGGACTCGACCCGCGGCAGCGAC 650 EAFESAGITPDSTRGSD ACCGGGGTGTTCGTCGGCGCCTTCTCCTACGGTTACGGCACCGGTGCGGA 700 TGVFVGAFSYGYGTGAD 50 CACCGACGGCTTCGGCGGCGACCGGCTCGCAGACCAGTGTGCTCTCCGGCC 750 T D G F G A T G S Q T S V L S G GGCTGTCGTACTTCTACGGTCTGGAGGGTCCGGCGGTCACGGTCGACACG 800 RLSYFYGLEGPAVTVDT GCGTGTTCGTCGTCGCTGGTGGCCCTGCACCAGGCCGGGCAGTCGCTGCG 850 55 ACSSSLVALHQAGQSLR CTCCGGCGAATGCTCGCTCGCCCTGGTCGGCGGCGTCACGGTGATGGCGT 900 S G E C S I. A L V G G V T V M A CTCCCGGCGGCTTCGTGGAGTTCTCCCGGCAGCGCGGCCTCGCGCCGGAC 950 SPGGFVEFSRQRGLAPD GGCCGGGCGAAGGCGTTCGGCGCGGGGTGCGGACGGCACGAGCTTCGCCGA 1000 60

	G R A K A F G A G A D G T S F A E GGGTGCCGGTGTGCTGATCGTCGAGAGGCTCTCCGACGCCGAACGCAACG	1050
5	G. A G V L I V E R L S D A E R N GTCACACCGTCCTGGCGGTCGTCGTGGTTCGGCGGTCAACCAGGATGGT G H T V L A V V R G S A V N Q D G	1100
5	GCCTCCAACGGGCTGTCGGCGCCGAACGGGCCGTCGCAGGAGCGGGTGAT	1150
	A S N G L S A P N G P S Q E R V I CCGGCAGGCCTGGCCAACGCCGGGCTGACGCGGGCGGACGTGGACGCCG R Q A L A N A G L T P A D V D A	1200
10	TOGAGGCCCACGGCACCGGCACCAGGCTGGGCGACCCCATCGAGGCACAG	1250
	GCGGTACTGGCCACCTACGGACAGGAGCGCGCCACCCCCTGCTGCTGGGACACCACCCCCCTGCTGCTGGGACACCACCCCCCTGCTGCTGGGACACCACCCCCCTGCTGCTGGGACACCACCCCCTGCTGCTGGGACACCACCACCACCACCACCACCACCACCACCACCAC	1300
15	CTCGCTGAAGTCCAACATCGGCCACGCCCAGGCCGCGTCCGGCGG S L K S N I G H A Q A A S G V A	1350
	GCATCATCAAGATGGTGCAGGGCCCTCCGGCACGGGGGGGG	1400
	CTGCACGCCGACGACGCCGCCGCACGTCGACTGGACGGCCGGC	1450
20	CGAACTGCTGACGTCGGCCGGCCGTGGCCCGAGACCGACC	1500
	GTGCCGCCGTCTCCTCGTTCGGGGTGAGCGGCACCAACGCCCACGTCATC R A A V S S F G V S G T N A H V I	1550
25	CTGGAGGCCGGACCGGTAACGGAGACGCCCGCGCATCGCCTTCCGGTGA L E A G P V T E T P A A 5 P S G D	1600
	CCTTCCCCTGCTGGTCGCCACGCTCACCGGAAGCGCTCGACGAGCAGA L P L L V S A R S P E A L D E Q	1650
	TCCGCCGACTGCGCGCCTACCTGGACACCACCCCGGACGTCGACCGGGTG I R R L R A Y L D T T P D V D R V	1700
30	GCCGTGGCACAGACGCTGGCCCGGCGCACACACTTCGCCCACCGCGCCGT A V A Q T L A R R T H F A H R A V	1750
	GCTGCTCGGTGACACCGTCATCACCACACCCCCGCGGACCGGCCCGACG L L G D T V I T T P P A D R P D	1800
35		1850
	GAGCAGCTAGCCGATTCGTCGGTGTGTTCGCCGAGCGGATGGCCGAGTG E Q L A D S S V V F A E R M A E C	1900
	TGCGGCGGCGTTGCGCGAGTTCGTGGACTGGGATCTGTTCACGGTTCTGG A A A L R E F V D W D L F T V L	1950
40	ATGATCCGGCGGTGGTGGACCGGGTTGATGTGGTCCAGCCCGCTTCCTGG D D P A V V D R V D V V Q P A S W	2000
	GCGATGATGGTTTCCCTGGCCGCGGTGTGGCAGGCGGCCGGTGTGCGGCC A M M V S L A A V W Q A A G V R P	2050
45	GGATGCGGTGATCGCCAGGGTGAGATCGCCGCAGCTTGTGTGG D A V I G H S Q G E I A A A C V	2100
	CGGGTGCGGTGTCACTACGCGATGCCGCCCGGATCGTGACCTTGCGCAGC A G A V S L R D A A R I V T L R S	2150
	CAGGCGATCGCCCGGGGCCTGGCGGGCCGGGGCGGGCGATGGCATCCGTCGC Q A I A R G L A G R G A M A S V A	2200
50	CCTGCCGCGCAGGATGTCGAGGGGCCTGGATCGCCGCCC L P A Q D V E L V D G A W I A A	2250
	ACAACGGGCCCGCCTCACCGTGATCGCGGGCACCCCGGAAGCGGTCGAC H N G P A S T V I A G T P E A V D	2300
55	CATGTCCTCACCGCTCATGAGGCACAAGGGGTGCGGGTGCGGCGGATCAC H V L T A H E A Q G V R V R R I T	2350
	CGTCGACTATGCCTCGCACACCCCGCACGTCGAGCTGATCCGCGACGAAC V D Y A S H T P H V E L I R D E	2400
	TACTCGACATCACTAGCGACAGCAGCTCGCAGACCCCGCTCGTGCCGTGG	2450
60	CTGTCGACCGTGGACGGCACCTGGGTCGACGGGGAGTA L S T V D G T W V D S P L D G E Y	2500
	CTGGTACCGGAACCTGCGTGAACCGGTCGGTTTCCACCCCGCCGTCAGCC W Y R N L R E P V G F H P A V S	2550
	AGTTGCAGGCCCAGGGCGACACCGTGTTCGTCGAGGTCAGCGCCAGCCCG	2600

	Q L Q A Q G D T V F V E V S A S P STGTTGTTGCAGGCGATGACGACGATGTCACGGTTGCCACGCTGCG V L L Q A M D D D V V T V A T L R	2650
5	V L L Q A M D D D V V T V A T L R TOGTGACGACGGCGACGCCACCGGGATGCTCACCGCCTGGCACAGGCCT R D D G D A T R M L T A L A Q A	2700
	ATGTCCACGGCGTCACCGTCGACTGGCCCGCCATCCTCGGCACCACCACA Y V H G V T V D W P A I L G T T T	2750
1.0	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	2800
10	L E S A R P A A S D A G H P V L	2850
	GCTCCGGTATCGCCCTCGCCGGGTCGCCGGGCCGGGTGTTCACGGGTTCC G S G I A L A G S P G R V F T G S GTGCCGACCGGTGCGGACCGCGGGTGTTCGTCGCCGAGCTGGCGCTGGC	
15	V P T G A D R A V F V A E L A L A CGCCGCGGGACGGGTCGACTGGGCCACGGTCGAGGGTCGACATCGCCT	3000
	A A D A V D C A T V E R L D I A COGTGCCGGCCGGCCGGCCGGCCATGGCCCGACGGTACAGACCTGGGTC	3050
20	S V P R P G H G R T T V Q T W V SACGAGCCGGCGACGCGCGCGCGCGCGCTTCACCCTGCACACCCGCAC D E P A D D G R R R F T V H T R T	3100
	CGGCGACGCCCGTGGACGCTGCACGCGGAGGGGGGGGGCCCCATG G D A P W T L H A E G V L R P H	3150
25	GCACGGCCTGCCCGATGCGGCCGACGCCGAGTGGCCCCACGGGCGCGGGCGG	
	GTGCCGGGACGGGCTGCCGGGTGTGTGGCGCGGGGGACCAGGTCTT V P A D G L P G V W R R G D Q V F	3250
30	CGCCGAGGCCGAGGTGGACGGACGGACGGTTTCGTGGTGCACCCCGACC A E A E V D G P D G F V V H P D TGCTUGACGCGGTCTTCTCCGCGGTCGGCGACGGAAGCCGCCAGCCGGCC	3300
	L L D A V F S A V G D G S R Q P A GGATGGCGCGACCGTACTGCGCGC	3400
0.5	G W R D L T V H A S D A T V L R A CTGCCTCACCCGGCGCACCGACGGAGCCATGGGATTCGCCGCCTTCGACG	3450
35	C L T R R T D G A M G F A A F D GCGCCGGCCTGCCGGTACTCACCGCGGAGGCGGTGACGCTGCGGGAGGTG	3500
	G A G L P V L T A E A V T L R E V GCGTCACCGTCCGAGGAGTCGGACGGCCTGCACCGGTTGGAGTG A S P S G S E E S D G L H R L E W	3550
40	GCTCGCGSTCGCCGAGGGGGTCTACGACGGTGACCTGCCCGAGGGACATG L A V A E A V Y D G D L P E G H	3600
	TCCTGATCACCGCCGCCCCCGACGACCCCGAGGACATACCCACCC	
45	GCCCACACCCGCGCCACCGGCGTCCTGACCGCCCTGCAACACCACCTCAC A H T R A T R V L T A L Q H H L T CACCACCGACCACCACCACCACCGCCGGCGGCG	
	T T D H T L I V H T T T D P A G CCACCGTCACCGCCCCACCGCCCCACCGC	
50	A T V T G L T R T A Q N E H P H R ATCCGCCTCATCGAAACCGACCCCCCCACACCCCCCTCCCCCTGGCCCA	
	I R L I E T D H P H T P L P L A Q ACTCGCCACCTCGACCACCCCCACCTCGCCTCACCCACCACACACCACCACCACCACCACCACCACCA	3900
55	L A T L D H P H L R L T H H T L ACCACCCCCACCTCCACCCCCCCCCCCCCCCCCCCCC	3950
23	CCCCTCAACCCCGAACACGCCATCATCATCACCGGCGGCTCCGGCACCCT PLNPEHAIIITGGSSGTL	4000
	CGCCGGCATCCTCGCCCGCCACCTGAACCACCCCCACACCTACCT	
60	CCCGCACCCCCCGACGCCACCCCGGCACCCACCTCCCCTGCGAC S R T P P P D A T P G T H L P C D	
	GTCGGCGACCCCACCAACTCGCCACCACCCTCACCCACATCCCCCAACC V G D P H Q L A T T L T H I P Q P CCTCACCGCCATCTTCCACACCGCCGCCACCCTCGACGACGGCATCCTCC	
	OFFINAL DEPONDED TO THE PROPERTY OF THE POPULATION OF THE POPULATI	1200

LTAIFHTAATLDDGIL ACGECETOACCCCGGACCGCCTCAC H A, L T P D R L T T V L H P K A N GCCGCCTGGCACCTGACCCAAAACCAACCCTCACCCACTT 4300 AAWELHELTONOPLTHE CGTCCTCTACTCCAGCGCCGCCGCCGTCCTCGGCAGCCCCGGGAGAAGGGAA 4350 V L Y S S A A A V L G S P G Q G N Y A A A N A F L D A L A T H R E 10 ADDOTOGGOCAACCOGCCAUCTCCATCGCCTGGGGCATGTGGCAEACCAE 4450 T 5 G Q P A T S I A W G M W H T 1 CAGCACCCTCACCGGACAACTCGACGACGCCGACGGGACCGCATCCGCC 4500 STLTGQLDDADRDRIE GCGGCGGTTTCCTCCCGATCACGGACGACGAGGGCATGGGGATGCAT 15 R G G F L P I T D D E G

20

25

30

35

Phage KC515 DNA was prepared using the Educe described in Genetic Manipulation of *Streptomyces*. A Laboratory Manual, edited by D. Hopwood *et al.* A phage suspension prepared from 10 plates (100 mm) of confluent plaques of KC515 on *S. lividans* TK24 generally gave about 3 µg of phage DNA. The DNA was ligated to circularize at the cos site, subsequently digested with restriction enzymes *Bam*HI and *Pst*I, and dephosphorylated with SAP.

Each module 8 cassette described above was excised with restriction enzymes Bg/III and NsiI and ligated into the compatible BamHI and PsiI sites of KC515 phage DNA prepared as described above. The ligation mixture containing KC515 and various cassettes was transfected into protoplasts of Streptomyces lividans TK24 using the procedure described in Genetic Manipulation of Streptomyces, A Laboratory Manual edited by D. Hopwood et al. and overlaid with TK24 spores. After 16-24 hr, the plaques were restreaked on plates overlaid with TK24 spores. Single plaques were picked and resuspended in 200 µL of nutrient broth. Phage DNA was prepared by the boiling method (Hopwood et al., supra). The PCR with primers spanning the left and right boundaries of the recombinant phage was used to verify the correct phage had been isolated. In most cases, at least 80% of the plaques contained the expected insert. To confirm the presence of the resistance marker (thiostrepton), a spot test is used, as described in Lomovskaya et al. (1997), in which a plate with spots of phage is overlaid with mixture of spores of TK24 and phiC31 TK24 lysogen. After overnight incubation, the plate is overlaid with antibiotic in soft agar. A working stock is made of all phage containing desired constructs.

Streptomyces hygroscopicus ATCC 14891 (see US Patent No. 3,244,592, issued 5 Apr 1966, incorporated herein by reference) mycelia were infected with the recombinant phage by mixing the spores and phage (1 x 10⁸ of each), and incubating on R2YE agar (Genetic Manipulation of Streptomyces, A Laboratory Manual, edited by D.

Hopwood et al.) at 30°C for 10 days. Recombinant clones were selected and plated on minimal medium containing thiostrepton (50 µg/mi) to select for the thiostrepton resistance-conferring gene. Primary thiostrepton resistant clones were isolated and purified through a second round of single colony isolation, as necessary. To obtain thiostrepton-sensitive revertants that underwent a second recombination event to evict the phage genome, primary recombinants were propagated in liquid media for two to three days in the absence of thiostrepton and then spread on agar medium without thiostrepton to obtain spores. Spores were plated to obtain about 50 colonies per plate, and thiostrepton sensitive colonies were identified by replica plating onto thiostrepton containing agar medium. The "CP was used to determine which of the thiostrepton sensitive colonies reverted to the wild type (reversal of the initial integration event), and which contain the desired AT swap at module 8 in the ATCC 14891-derived cells. The PCR primers used amplified either the KS/AT junction or the AT/DH junction of the wild-type and the desired recombinant strains. Fermentation of the recombinant strains. followed by isolation of the metabolites and analysis by LCMS, and NMR is used to characterize the novel polyketide compounds.

5

10

15

20

25

30

35

Example 2

Replacement of Methoxyl with Hydrogen or Methyl at C-13 of FK-506

The present invention also provides the 13-desmethoxy derivatives of FK-506 and the novel PKS enzymes that produce them. A variety of *Streptomyces* strains that produce FK-506 are known in the art, including *S. tsukubaensis* No. 9993 (FERM BP-927), described in U.S. Patent No. 5,624,852, incorporated herein by reference; *S. hygroscopicus* subsp. *yakushimaensis* No. 7238, described in U.S. patent No. 4,894,366, incorporated herein by reference; *S.* sp. MA6858 (ATCC 55098), described in U.S. Patent Nos. 5,116,756, incorporated herein by reference; and *S.* sp. MA 6548, described in Motamedi *et al.*, 1998, "The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK-506," *Eur. J. Biochem. 256*: 528-534, and Motamedi *et al.*, 1997, "Structural organization of a multifunctional polyketide synthase involved in the biosynthesis of the macrolide immunosuppressant FK-506," *Eur. J. Biochem. 244*: 74-80, each of which is incorporated herein by reference.

The complete sequence of the FK-506 gene cluster from *Streptomyces* sp. MA6548 is known, and the sequences of the corresponding gene clusters from other FK-506-producing organisms is highly homologous thereto. The novel FK-506 recombinant gene clusters of the present invention differ from the naturally occurring gene clusters in

that the AT domain of module 8 of the naturally occurring PKSs is replaced by an AT domain specific for malonyl CoA or methylmalonyl CoA. These AT domain replacements are made at the DNA level, following the methodology described in Example 1.

The naturally occurring module 8 sequence for the MA6548 strain is shown below, followed by the illustrative hybrid module 8 sequences for the MA6548 strains.

5

	GCATGCGGCTGTACGAGGCGCACGGGCACCGGAAGTCCCGTG3TG3TG 5	С
10	GCGGCCCCCTCGACGACGCCCGACGTGCCGCTGCTGCGCGGGGCTGCC	100
10	000100000	150
	R T T V R R A A V R E R S L A D GUTGGCGTGCTGCCGACGACGACGACGCGCGACGACGACGACGACGACGAC	200
15	TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCGAAGACAT SWNSTATVLGHLGAEDI	250
	CCCGGCGACGACGTTCAAGGAACTCGGCATCGACTCGCTCACCGCGG	300
20	TCCAGCTGCGCAACGCGTGACCACGGCGGACCGGCGTACGCCTGAACGCC	350
	ACAGOGGTCTTCGACTTTCCGACGCCGCGCGCGCGCGCGCGCGAGACTCGG T A V F D F P T P R A L A A R L G	400
	CGACGAGCTGGCCGGTACCCGCGCGCGCCGGTCGCGGGCCGGGCCGGGCCACCGGGCCCGCGCGCG	450
25	CCGCGGCCGCACGACGAACCGCTGGCGATCGTGGGCATGGCCTGCCGT T A A A H D E P L A I V G M A C R	500
	CTGCCGGGCGGGTCGCGTCGCCACAGGAGCTGTGGCGTCTCGCGTCGCGTCL PGGVASPQELWRLVAS	550
30	CGGCACCGACGCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG G T D A I T E F P A D R G W D V	
	ACGCGCTCTACGACCCGGACCCCGACGCGATCGGCAAGACCTTCGTCCGG D A L Y D P D P D A I G K T F V R	
2.5	H G G F L D G A T G F D A A F F G	700
35	I S P R E A L A M D P Q Q R V L	750
	TGGAGACGTCCTGGGAGGCGTTCGAAAGCGCGGGCATCACCCCGGACGCG L E T S W E A F E S A G I T P D A GCGCGGGGCAGCGACACCGGCGTGTTCATCGGCGCGTTCTCCTACGGGTA	
40	A R G S D T G V F I G A F S Y G Y CGGCACGCGTGCGGATACCAACGGCTTCGGCGCGACAGGGTCGCAGACCA	
	G T G A D T N G F G A T G S Q T GCGTGCTCTCCGGCCCCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG	
45	S V L S G R L S Y F Y G L L G P S GTCACGGTCGACACCGCCTGCTCGTCGTCACTGGTCGCCCTGCACCAGGC	1000
	V T V D T A C S S S L V A L H Q A AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTCGGCGGTG	1050
	G Q S L R S G E C S L A L V G G TCACGGTGATGGCGTCGCCCGGCGGCAGTTCGTCGAGTTCTCCCGGCAGCGC	1100
50	V T V M A S P G G F V E F S R Q R GGGCTCGCGCCGGACGGGCGGGCGAAGGCGTTCGGCGCGGGCGG	1150
	G L A P D G R A K A F G A G A D G TACGASCTTCGCCGAGGGCGCCGGTGCCCTGGTGGTCGAGCGGCTCTCCG	1200
55	T S F A E G A G A L V V E R L S ACGCGGAGCGCCACGGCCACGCGTCCTCGCCCTCGTACGCGGCTCCGCG	1250
	DAERHGHTVLALVRGSA GCTAACTCCGACGCGCGCGCGCGCGCCCCCC ANSDGASNGLSAPNGPS	1300
	CCAGGAACGCGTCATCCACCAGGCCTCGCGAACGCGAAACTCACCCCG	1350

Q E R V I H Q A L A N A K L T P CCGATGTCGACGCGGTCGACGCGCACGGCACCCGGCCTCGGCGAC 1400 A D V D A V E A H G T G T R L G D COATOBAGGOGTAGGUGGTGUTGGGGAGGTAGGGAGAGGAGGGGGGAG 1450 5 PIEAQALLATY3QDRA GCCCCTCCTGCTCGGCTGGCTGAAGTCGAACATCGGCGCACGCCCAGGCCG 1500 P L L L G S L K S N I G H A Q A CGTCAGGGGTCGCGGGATCATCARGATGGTGCAGGCCATCCGGCACGGG 1550 ASGVAGIIKKVQAIRHG 10 GAACTGCCGCCGACACTGCACGCGGACGAGCCGTCGCCGCACGTCGACTG 1600 ELPPTLHADEPSPHVDW T A G A V E L L T S A R P W P G COGGTCCCCCGCGCGCGCGCCGTCCCTCGTCGTTCGGCGTGAGCGCACG 1700 15 T G R P R R A A V S S F G V S G T AACGCCCACATCATCCTTGAGGCAGGACCGGTCAAAACGGGACCGGTCGA 1750 NAHIILEAGPVKTGPVE GGCAGGAGCGATCGAGGCAUGACCGGTCGAAGTAGGACCUGTCGAGCCTG 1800 A G A I E A S P V E V G P V E A GACCSCTCCCCGCGGGGCGCCGCCGTCAGCACCGGGGGAAGACCTTCCGCTG 1850 20 G P L P A A P P S A P G E D L P I L V S A R S P E A L D E Q I G R I 25 RAYLOTGPSVORAAVA AGACACTGGCCGGGGTAGGCACTTCAGCCACGGGGCGTACTGCTCGGG 2000 Q T L A R R T H F T H R A V L L G GACACCGTCATCGGCGCTCCCCCCGCGGACCAGGCCGACGAACTCGTCTT 2050 D T V I G A P P A D Q A D E L V F 30 CGTCTACTCCGGTCAGGGCACCCAGCATCCCGCGATGGGCGAGCAACTCG 2100 V Y S G Q G T Q H P A M G E Q L CGGCCGCGTTCCCCGTGTTCGCCGATGCCTGGCACGCGCGCCCCGACGG 2150 A A A F P V F A D A W H D A L R R CTCGACGACCCGGACCGGCACGACCCCACACGGAGCCAGCACACGCTCTT 2200 35 L D D P D P H D P T R S Q H T L F CGCCCACCAGGGGGCGTTCACCGCCCTCCTGAGGTCCTGGGACATCACGC 2250 A H Q A A F T A L L R S W D I T CGCACGCCGTCATCGGCCACTCGCTCGGCGAGATCACCGCCGCGTACGCC 2300 P H A V I G H S L G E I T A A Y A GCCGGGATCCTGTCGCTCGACGACGCCTGCACCCTGATCACCACGCGTGC 2350 . 40 A G I L S L D D A C T L I T T R A CCGCCTCATGCACACGCTTCCGCCGCCCGCCGCCCATGGTCACCGTGCTGA 2400 R L M H T L P P P G A M V T V L CCAGCGAGGAGGAGGCCCGTCAGGCGCTGCGGCCGGGGGGTGGAGATCGCC 2450 45 T S E E E A R Q A L R P G V E I A GCGGTCTTCGGCCCGCACTCCGTCGTGCTCTCGGGCGACC13GACGCCGT 2500 A V F G P H S V V L S G D E D A V GCTCGACGTCGCACAGCGGCTCGGCATCCACCACCGTCTGCCCGCGCCGC 2550 L D V A Q R L G I H H R L P A P 50 ACGCGGGCCACTCCGCGCACATGGAACCCGTGGCCGCCGAGCTGCTCGCC 2600 H A G H S A H M E P V A A E L L A ACCACTCGCGAGCTCCGTTACGACCGGCCCACACCGCCATCCCGAACGA 2650 T T R E L R Y D R P H T A I P N D CCCCACCACCGCCGAGTACTGGGCCGAGCAGGTCCGCAACCCCGTGCTGT 2700 55 P T T A E Y W A E Q V R N P V L TCCACGCCCACACCCAGCGGTACCCCGACGCGTGTTCGTCGAGATCGGC 2750 F H A H T Q R Y P D A V F V E I G CCCGGCCAGGACCTCTCACCGCTGGTCGACGGCATCGCCCTGCAGAACGG 2800 PGQDLSPLVDGIALQNG 60 TADEVHALHTALARLF CACGCGGCGCCACGCTCGACTGGTCCCGCATCCTCGGCGGTGCTTCGCGG 2900 TRGATLDWSRILGGASR CACGACCCTGACGTCCCTCGTACGCGTTCCAGCGCGTCCCTACTGGAT 2950

	H D	_											W I	
													TCGGCA	3000
	E S OCCGAG		Р						G				L G	2050
5)											3050
_	000000													3100
	P A												A A	
	CGCCGA									TCG	ACG:	FCA	CCTCCG	3150
10			T								5 7			2000
10													GTCGAT	3200
	V P GAACCO									_				3250
	E P) G					r v				V G	3230
	CGACGC	cccc	TGGA	CGC'	TGCA	agac	:GAG	:GGG:	GTTC	maa	GDCG	003	geeges	3300
15		. ?			L H								G R	
	TGCCCC													3350
	V P CCCGCG													3400
	P A												F V	2400
20	CGAAGC													3450
	E A	E	V	D :	S P	Ð	G	E	Λ	Ã	H I	?	D L	
													ACCGGA	3500
			/ F										T G	2550
25	TGGCGC W R				JUADO H								A C	3550
	CCTCAC													3600
	L T				S G								D G	
	CCGGAA		CGGTG	GCTC										3650
30	A G		o V						T					3700
30	TCGGCA S A												W L	2/00
	GCCGGT													3750
	P V				H Y				D					
25	ACACCC													3800
35		LI		A	T				P			P		2050
	CCCCAC P H	AACE N		CAC T		KUUU T	H		haac 2 T				L T	3850
	-					_			-				TOCACA	3900
	A L	Q	Н	H I	L I	Τ	Ţ	N	Н	T	L		√ H	
4 ()													ACCGCA	3950
	TT												T A CCCCCA	4000
													P H	4000
													ACCTAC	4050
45	T P	L	P	L '	r ç	L	T	~	Ĭ.	Н	Q E)	H L	
													ACCACC	4100
	R L												T T CCCCAA	41E0
													P N	4130
50													TCCTCG	4200
	A H	I	L	Ι '	T G	G	S	G	T	L	A (3	I L	
													CCACCA	4250
	A R												P P CCCCAC	4300
55													P T	4300
	CCAAAT													4350
	QI	Т	Q	A	L T	Н	_	P	Q	P	L 7		G I	
													ACCCCC	4400
60	F H													4450
60													GCACCT H L	4450
													ACTCCA	4500
	H F	H H	T	Q	N Q	₽		T	H	F	V 1	Ĺ	Y S	
	GCGCCC												GCCGCC	4550

10

5

The AvrII-XhoI hybrid FK-506 PKS module 8 containing the AT domain of module 12 of rapamycin is shown below.

GCATGOGGCTGTACGAGGCGGCACGGCACCGGAAGTCCCGTGGTGGTG 50 M R L Y E A A R R T G S P V V V 15 DOGGEOGGGCTCGACGACGCCGGGCCGGACGTGCCGCTGCTGCGCGGGCTGCG 100 A A A L D D A P D V P L L L B GOGTA CGA COGT COGCOT GCC GCCGT CCGGGAA CGCT CT CCCCCACC - 1 8 1 R T T V R R A A V R E F S L A D 20 RSPCCPTTSAPTPPSRS TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCGAAGACAT 250 S W N S T A T V L G R L G A E D I CCCGGCGACGACGACGTTCAAGGAACTDGGCATCGACTCGCTCACCGCGG 300 PATTTFKELGICSLTA 25 TCCAGCTGCGCAACGCGCTGACCACGCGGCGACCGGCGTACGCCTCAACGCC 350 V Q L R N A L T T A T G T R L N A ACAGCGGTCTTCGACTTTCCGACGCCGCGCGCGCGCCCCCCGAGACTCGG 400 TAVFDFPTPRALAARLG CGACGAGCTGGCCGGTACCCGCGCGCCCGTCGCGGCCGGACCGCGGCCA 450 30 DELAGTRAPVAARTAA CCGCGGCCGCACGACGACCGCTGGCGATCGTGGCCATGCCTT 500 TAAAHDEPLAIVGMACR CTGCCGGGCGGGTCGCCTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550 L P G G V A S P O E L W R L V A S 35 CGGCACCGACGCCATCACGGAGTTCCCCGCGGGACCGCGGCTGGGACGTGG 600 G T D A I T E F P A D R G W D V ACGCGCTCTACGACCCGGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650 DALYDPDPDAIGKTFVR CACGGCGGCTTCCTCGACGGTGCGACCGGCTTCGACGCGCGTTCTTCGG 700 40 H G G F L D G A T G F D A A F F G GATCAGCCGCGCGAGGCCCTGGCCATGGACCCGCAGCAACGGGTGCTCC 750 I S P R E A L A M D P Q Q R V L TGGAGACGTCCTGGGAGGCGTTCGAAAGCGCGGGGCATCACCCCGGACGCG 800 L E T S W E A F E S A G I T P D A GCGCGGGGCAGCGACACCGGCGTGTTCATCGGCGCGTTCTCCTACGGGTA 850 ARGSDTGVFIGAFSYGY CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGAJAGGGTCGCAGACCA 900 G T G A D T N G F G A T G S Q T GCGTGCTCTCCGGCCGCCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950 50 S V L S G R L S Y F Y G L E G P S GTCACGGTCGACACCGCCTGCTCGTCGTCACTGGTCGCCCTGCACCAGGC 1000 V T V D T A C S S S L V A L H Q A AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTCGGCGGTG 1050 G Q S L R S G E C S L A L V G G TCACGGTGATGGCGTCGCCCGGCGGATTCGTCGAGTTCTCCCGGCAGCGC 1100 55 V T V M A S P G G F V E F S R Q R GGGCTCGCGCGGACGGGCGGGCGAAGGCGTTCGGCGCGGGCGCGGACGG 1150 G L A P D G R A K A F G A G A D G TACGAGCTTCGCCGAGGGCGCCGGTGCCCTGGTG3TCGAGCGGCTCTCCG 1200 60 T S F A E G A G A L V V E R L S ACGCGGAGCGCCACGGCCACACCGTCCTCGCCCTCGTACGCGGCTCCGCG 1250 DAERHGHTVLALVRGSA

GCTAACTCCGACGGGGGGTCGAACGGTCTGTCGGGGGGCGGAACGGCCCCTC 1300 A N S D G A S N G L S A P N G P S CCAGGAACGCGTCATCCACCAGGCCCTCGCGAACGCGAAACTCACCCCCG 1350 Q E R V I H Q A L A N A K L T P COGATGTOGACGCGGTOGAGGGGCACGGGAACGCGGCCTCGGCGAC 1400 A D V D A V E A H G T G T R L G D CCONTCONGCCCAMGGCCCTCCTCGCCNCGTACGGACACACACCACGGGGGGGAC PIEAQALLATYS Q D R A T GCCCCTGCTGCTCGGCTGAAGTCGAACATCGGGCAGGCCAGGCCAGGCCG 10 PLLLGSLKSNIGHAQA CGTCAGGGGTCGCCGGGATCATCAAGATGGTGCAGGCCATCCGGCACGG3 1550 A S G V A G I I K M V Q A I R H G GAACTGCCGCCGACACTGCACGCGGACGCGACGTCGCCGCACGTCGACTG 1600 E L P P T L H A D E P S P H V D W 15 GACGGCCGGTGGAGCTCCTGACGTCGGCCCGGCCGTGGCCGGGGA 1650 TAGAVELLTSARPWPG OCGGTCGCCCTACCCCTCAGGCGTGTCGTCCTTCGGGGATCAGTGGCACC 1700 TGRELLAGVSSFGISG. AACGCCCAUGTCATCCTGGAAAGCGCACCCCCACTCAGCCTGCGGACAA 1750 20 NAHVILESAPPTQPADN CGCGGTGATCGAGCGGGCACCGGAGTGGGTGCCGTTGGTGATTTCGGCCA 1800 A V I E R A P E W V P L V I S A RTQSALTEHEGRLRAYL 25 GCGCCTCGCCCGCCTGCATATGCGCGCCTGTGGCATCGACGCTGGCCAT 1900 A A S P G V D M R A V A S T L A M GACACGGTCGGTGTTCGAGCACCGTGCCGTGCTGCTGCTGGGAGATGACACCG 1980 T R S V F E H R A V L L G D D T TCACCGGCACCGCTGTGTCTGACCCTCGGGCGGTGTTCGTCTTCCCGGGA 2000 30 V T G T A V S D P R A V F V F P G CAGGGGTCGCAGCGTGCTGGCATGGGTGAGGAACTGGCCGCCGCGTTCCC 2050 Q G S Q R A G M G E E L A A A F P CGTCTTCGCGCGGATCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCG 2100 V F A R I H Q Q V W D L L D V P 35 ATCTGGAGGTGAACGAGACCGGTTACGCCCAGCCGGCCCTGTTCGCAATG 2150 D L E V N E T G Y A Q P A L F A M CAGGTGGCTCTGTTCGGGCTGCTGGAATCGTGGGGTGTACGACCGGACGC 2200 Q V A L F G L L E S W G V R P D A GGTGATCGGCCATTCGGTGGGTGAGCTTGCGGCTGCGTATGTGTCCGGGG 2250 40 V I G H S V G E L A A A Y V S G TGTGGTCGTTGGAGGATGCCTGCACTTTGGTGTCGGCGCGGGGCTCGTCTG 2300 V W S L E D A C T L V S A R A R L ATGCAGGCTCTGCCCGCGGGTGGGGTGATGGTCGCTGTCCCGGTCTCGGA 2350 M O A L P A G G V M V A V P V S E 45 GGATGAGGCCGGGCCGTGCTGGGTGAGGSTGTGGAGATCGCCGCGGTCA 2400 ACGGCCCGTCGTCGCTGCTCTCTCCGGTGATGAGGCCGCCGTGCTGCAG 2450 N G P S S V V L S G D E A A V L Q GCCGCGGAGGGGCTGGGAAGTGGACGCGGCTGGCAACCACCTACGCGTT 2500 50 A A E G L G K W T R L A T S H A F CCATTCCGCCCGTATGGAACCCATGCTGGAGGAGTTCCGGGCGGTCGCCG 2550 H S A R M E P M L E E F R A V A AAGGCCTGACCTACCGGACGCCGCAGGTCTCCATGGCCGTTGGTGATCAG 2600 EGLTYRTPQVSMAVGDQ 55 GTGACCACCGCTGAGTACTGGGTGCGGCAGGTCCGGGACACGGTCCGGTT 2650 V T T A E Y W V R O V R D T V R F CGGCGAGCAGGTGGCCTCGTACGAGGACGCCGTGTTCGTCGAGCTGGGTG 2700 G E Q V A S Y E D A V F V E L G CCGACCGGTCACTGGCCCGCCTGGTCGACGGTGTCGCGATGCTGCACGGC 2750 A D R S L A R L V D G V A M L H G 60 GACCACGAAATCCAGGCCGCGATCGGCGCCCTGGCCCACCTGTATGTCAA 2800 D H E I Q A A I G A L A H L Y V N CGGCGTCACGGTCGACTGGCCCGCGCTCCTGGGCGATGCTCCGGCAACAC 2850 $\mathsf{G} \quad \mathsf{V} \quad \mathsf{T} \quad \mathsf{V} \quad \mathsf{D} \quad \mathsf{W} \quad \mathsf{P} \quad \mathsf{A} \quad \mathsf{L} \quad \mathsf{L} \quad \mathsf{G} \quad \mathsf{D} \quad \mathsf{A} \quad \mathsf{P} \quad \mathsf{A} \quad \mathsf{T}$

	GGGTGCTGGACCTTCCGACATACGCCTTCCAGCACCAGCGCTACTGGCTC R V L D L P T Y A F Q E Q R Y W L	2900
	GAGTOSGOTOCCCGGGCCACCCGGGCCACCCCGGCACCCGGCACCCGGCACCCGGCACCCGGCACCCGGGCCACCCGGGCCACCCGGGCCACCCGGGCCACCCGGGCCACCCGGGCCACCCGGGCCACCGGGCACCGGGCACCGGGCACGGGCACGGGCACGGGCACGGGCACGGGCACGGGCACGGGCACGGGCACGGGCACGGACGGGCACGGACGGACGAC	2950
5	CSGASTCSCCGTCGCCGGGGTCGCCGGGGCGGGTGTTCACGGGTCCCGTGC	3000
	COCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	3050
10	GCCGACGCCACCGACTGCGCCACGGTCSAACAGCTCACGTCACCTCCGT A D A T D C A T V E O L D V T S V	3100
	GCCCGGCGGATCCGCGCGCGCGGGGGCACCGCGCGAGACCTGGGTCGATG P G C S A R G R A T A O T W V D	3150
	AACCCCCCCCACACGCCCCCCCCCCCCCCCCCCCCCCCC	3200
15	SACOCCCCCTGGACGCTGCACCCCGACGGGCCCCCGGCCGCCGCCGCCGCCGCCGCC	3250
	D A P W T L H A E G V L R P G R V GCCCCAGCCGGAAAGCCGTCGACACCGCCTGGCCCGGCGGGGGGGG	3300
	P Q P E A V D T A W P P P G A V COGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	3350
20	P A D G L P G A W R R A D Q V F 7 GAAGCCGAAGCCGACCCGACCTGACGCCTGCTGGCACACCCCGACCTGCT	3400
	EAEVDSPDGFVAHPDLL	3450
2.5	CGACGCGGTCTTCTCCGCGGTCGGCGACGGGAGCGGACCGGAT D A V F S A V G D G S R Q P T G	
25	GGCGCACCTCGCGGTCGCACGCCGCGCGCGCGCGCGCGCG	3500
	CTCACCCGCCCCGACAGTGGTGTCCTCGACCGTCGCCGCCTTCGACCGTGC L T R R D S G V V E L A A F D G A	3550
30	CGGAATGCCGSTGCTCACCGCGGAGTCGGTGACGCTGGGGGGGGGCGCGT	3600
	CGGCAGGCGGATCGGACGAGTCGGACGGTCTGCTTCGGCTTGAGTGGTTG S A G G S D E S D G L L R L E W L	3650
	CCGGTGGCGGAGGCCCACTACGACGGTGCCGACGAGCTGCCCGAGGGCTA	3700
35	P V A E A H Y D G A D E L P E G Y CACCETCATCACCGCCACACCCGACGACGACCCGACGACCCGACGACCCGACGA	3750
	T L I T A T H P D D P D D P T N CCCACACACACACACACACACACACACACACACACACA	3800
	PHNTPTRTHTQTTRVLT	3850
. 40	A L Q R H L I T T N R T L I V H T CACCACCACCACCCAGGCGCGCGCGCACCGCACCGCA	3900
	T T D P P G A A V T G L T R T A AAAACGAACACCCGGGCGGCATCCACCTCATCGAAACCCACCACCACCACCACCACCACCACCACCACCAC	3950
4.5	Q N E H P G R I H L I E T H H P H	
45	ACCCCACTCCCCCTCACCCACCTACCACCTACCACCTACCACC	
	COTCACCAACAACACCCTCCACACCCCCACCTCACCCCCATCACCAC	4050
50	ACCACAACACCACCACCACCACCCCCACCCCACCCCACCCC	4100
	CACGCCATCCTCATCACCGGCGGCTCCGGCACCCTCGCCGGCATCCTCGC H A I L I T G G S G T L A G I L A	4150
	CCGCCACCTCAACCACCCCACACCTACCTCCTCTCCCGCACACCACCACCACCACCACCACCACCACCACCACC	4200
55	CCCCCACCACACCCGGCACCCACATCCCCTGCGACCTCACCGACCCCACC	4250
	P P T T P G T H I P C D L T D P T CAAATCACCCAAGCCCTCACCCACACACACACACCCCTCACCGGCATCTT	4300
	Q I T Q A L T H I P Q P L T G I F CCACACGGCGCCACCCTCGACGACGCCCCCCCCC	4350
60	H T A A T L D D A T L T N L T P AACACCTCACCACCACCCTCCAACCCAAAGCCGACGCCGCCTGGCACCTC	4400
	Q H L T T T L Q P K A D A A W H L CACCACACACCCAAAACCAACCCTGACCCACTTCGTCCTCTACTCCAG	
	H H H T Q N Q P L T H F V L Y S S	770

The AvrII-XhoI hybrid FK-506 PKS module 8 containing the AT domain of module 13 of rapamycin is shown below.

GCATGCGGCTGTACGAGGCGGCACGGCACCGGAAGTCCCGTGGTGGTG 15 MRLYEAARRTGSPVVV GOGSTOGOGOTOGACGACGCGCGGACGTGCCGCTGCTTCTGCGGGGCTGCG-100 A A A L D D A P D V P L L R G L R RTTVRRAAVRERSLAD 20 RSPCCPTTSAPTPPSRS TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCCCGAAGACAT 250 S W N S T A T V L G H L G A E D I CCCGGCGACGACGTTCAAGGAACTCGGCATCGACTCGCTCACCGCGG 300 25 PATTTFKELGIDSLTA TOCAGOTGOGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGCC 350 V Q L R N A L T T A T G V R L N A ACAGCGGTCTTCGACTTTCCGACGCCGCGCGCGCTCGCCGCGAGACTCGG 400 TAVFDFPTPRALAARLG 30 CGACGAGCTGGCCGGTACCCGCGCGCCCGTCGCGGCCCGGACCGCGGCCA 450 DELAGTRAPVAARTAA CCGCGGCCGCACGACGACGCTGGCGATCGTGGGCATGGCCTGCCGT 500 T A A A H D E P L A I V G M A C R CTGCCGGGGGGGGTCGCGTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550 35 L P G G V A S P O E L W R L V A S CGGCACCGACGCCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG 600 GTDAITEFPADRG W D V ACGCGCTCTACGACCCGGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650 DALYDPDPDAIGKTFVR 40 CACGGCGGCTTCCTCGACGGTGCGACCGGCTTCGACGCGGCGTTCTTCGG 700 H G G F L D G A T G F D A A F F G GATCAGCCCGCGCGAGGCCCTGGCCATGGACCCGCAGCAACGGGTGCTCC 750 I S P R E A L A M D P Q Q R V L TGGAGACGTCCTGGGAGGCGTTCGAAAGCGCGGGCATCACCCCGGACGCG 800 45 L E T S W E A F E S A G I T P D A GCGCGGGGCAGCGACACCGGCGTGTTCATCGGCGGGTTCTCCTACGGGTA 850 ARGSDTGVFIGAFSYGY CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTCGCAGACCA 900 G T G A D T N G F G A T G S Q T 50 GCGTGCTCTCCGGCCGCCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950 S 7 L S G R L S Y F Y G L E G P S GTCACGGTCGACACCGCCTGCTCGTCGTCACTGGTCGCCCTGCACCAGGC 1000 V T V D T A C S S S L V A L H Q A AGGGAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTCGGCGGTG 1050 55 G O S L R S G E C S L A L V G G TCACGGTGATGGCGTCGCCGGGGGGATTCGTCGAGTTCTCCCGGGCAGCGC 1100 V T V M A S P G G F V E F S R Q R GGSCTCGCGCCGGACGGGCGGGCGGAAGGCGTTCGGCGCGGGGCGGGACGG 1150 G L A P D G R A K A F G A G A D G TACGAGCTTCGCCGAGGGCGCCGGTGCCCTGGTGGTCGAGCGGCTCTCCG 1200 60 T S F A E G A G A L V V E R L S ACGCGGAGCGCCACGGCCACACCGTCCTCGCCCTCGTACGCGGCTCCGCG 1250

	D A E R H G H T V L A L V R G S A GCTAACTCCGACGGCGCTCGAACGGTCTGTCGGCGCCGAACGGCCCCTC	1300
5	A, N S D G A S N G L S A P N G P S CDAGGAACGCGTCATCCACCAGGCCCTCSCGAACGCSAACCCCCCG Q E R V I H Q A L A N A K L T P	1350
		1400
	COCATOGAGGOGCAGGOGCTGCTCGCGACGACGACAGGACAG	1450
10	OCCOURGE GOT CONTRACTOR OF A CATCOR OF A C	1500
	CGTCAGGGGTCGCCGGGATCATCAAGATGGTGCAGGCCATCCGGGCACGGG A S G V A G I I K M V Q A I R H G	1550
15	GAACTGCGGCGACACTGCACGCGGACGACGTCGCCGCACGTCGACTG E L P P T L H A D E P S P H V D W	1600
	T A G A V E L L T S A R P W F G	1850
	COGGTCGCCCTAGGCGGGGGGGGGGGGTGTCGTCCTTCGGAGTCAGCGCGACC T G R P R R A G V S S F G S G T	1700
20	AACGCCCACGICATCCTGGAGAGCGCACCCCCGGCTCAJUCGGGGAGGA N A H V I L E S A P F A Q F A E E	1750
	GGGGCAGCCTGTTGAGACGCCGGTGGTGGCCTGGGATGTGCCGCTGG A Q P V E T P V V A S D V L P L	1800
25	V I S A K T Q P A L T E H E D R L	1850
	RAYLAAS PGADIRAVAS	1900
20	T L A V T R S V F E H R A V L L	1950
30	GAGATGACACCGTCACCGGCACCGCGGTGACCGACCCCAGGATCGTGTTT G D D T V T G T A V T D F R I V F	
	GTCTTTCCCGGGCAGGGGTGGCAGTGGCAGTGGGCAGTGCACTGCG V F P G Q G W Q W L G M G S A L R	2050
35	CGATTCGTCGGTGGTGTTCGCCGAGCGGATGGCCGAGTGTGCGGCGCGT D S S V V F A E R M A E C A A A	2100
	TGCGCGAGTTCGTGGACTGGGATCTGTTCACGGTTCTGGATGATCCGGCG L R E F V D W D L F T V L D D P A	
40	GTGGTGGACCGGGTTGATGTGGTCCAGCCCCCTTCCTGGGCGATGATGGT V V D R V D V V Q P A S W A M M V	2200
40	TTCCCTGGCCGCGGTGTGGCAGGCGGGCGGGTGAGGCGGTGAGGCGGTGAGGCGCATTCGCAGGGTGAGATCGCGCAGCTTGTGTGGGGGGGTGCGGTG	
	I G H S Q G E I A A A C V A G A V TCACTACGCGATGCGGGATCGTGACCTTGCGCAGCCAGGCGATCGC	
45	S L R D A A R I V T L R S Q A I A CCGGGGGCCTGGCGGGGGGGGGGGGGGGGGGGGGGGG	
	R G L A G R G A M A S V A L P A AGGATGTCGAGGTGGTCGACGGGGCCTGSATCGCCGCCCACAACGGGCCC	
50	Q D V E L V D G A W I A A H N G P GCCTCCACCGTGATCGCGGGCACCCCGGAAGCGGTCGACCATGTCCTCAC	
	A S T V I A G T P E A V D H V L T CGCTCATGAGGCACAAGGGGTGCGGGTGCGGCGCATCACCGTCGACTATG	
	A H E A Q G V R V R R I T V D Y CCTCGCACACCCCGCACGTCGAGCTGATCCGCGACGACTACTCGACATC	2600
55	A S H T P H V E L I R D E L L D I ACTAGCGACAGCTCGCAGACCCCGCTCGTGCCGTGGCTGTCGACCGT	2650
	T S D S S S Q T P L V P W L S T V GGACGGCACCTGGGTCGACAGCCCGCTGGACGGGGAGTACTGGTACCGGA	2700
60	D G T W V D S P L D G E Y W Y R	
	ACCTGCGTGAACCGGTCGGTTTCCACCCCCCCGTCAGCCAGTTGCAGGCC	2750
	- -	

AMDDDVVTVATLRRDD GOGACGCCACCCGGATGCTCACCGCCCTGGCACGGCCTATGTCCACGGC 2900 G D A T R M L T A L A , A Y V H G STCACOGTOBACTOBCCOGCCATCCTCGBCACCACAACCCGGGTACT 2950 V T V D W F A I L G T T T R V L GGACCTTCCGACCTACGCCTTCCAACACCACCGGCGTACTGGGTCGGAGTCGG 3000 D L P T Y A F Q H Q R Y W L E S OTOCOCCACACCACCANCTCAGGCCACCCCCTCCTCGGCACCCGAAGTC 3050 A P P A T A D S G H P U L G T G V 10 GCCGTCGCCGGGTCGCCGGGGCCGGGTTTTCACGGGTCCCGTGCCCGCCGG 3100 A V A G S P G R V F T G P V P A G A D R A V F I A E L A L A A A D CCACCGACIGOGCCACGGTCGAACAGCTCGACGTCACCTCCGCTGCCCGGC 3200 15 A T D C A T V E Q L D V T S V P G GGATCCGCCGCGCAGGGCCACCGCGCAGACCTGGGTCGATGAACCCGC 32E0 S S A R G R A T A Q T W V D E P A CGCCGACGGGGGAAGGCGCTT MCCGTGCAUACCGGGGGGGGGGGGCGCC 3366 A D G .. R R F T V H T R V G D A 20 PWTLHAEGVLRP3RVPO PEAVDTAWPPPGAVPAD CGGGCTGCCGGGGGGTGGCGACGCGGGGGCCAGGTCTTCGTCGAAGCCG 3450 25 G L P G A W R R A D Q V F V E A AAGTEGACAGCCTTGACGGCTTCGTESCACACCCCEGACGTGCTCGACGCG 3500 E V D S P D G F V A H P D L L D A GTCTTOTOCGCGGTCGGCGACGGGAGCCGCCAGCCGATGGCGCGA 3550 V F S A V G D G S R Q F T G W R D 30 L A V H A S D A T V L R A C L T GCCGCGACAGTGGTGTCGTGGAGCTCGCCGCCTTCGACGGTGCCGGAATG 3650 R R D S G V V E L A A F D G A G M CCGGTGCTCACCGCGGAGTCGGTGACGCTGGGCGAGGTCGCGTCGGCAGG 3700 35 P V L T A E S V T L G E V A S A G CGGATCCGACGAGTCGGACGGTCTGCTTCGGCTTGAGTGGTTGCCGGTGG 3750 G S D E S D G L L R L E W L P V CGGAGGCCCACTACGACGGTGCCGAGGAGCTGCCCGAGGGCTACACCCTC 3800 A E A H Y D G A D E L P E G Y T L ATCACCGCCACACACCCCGACGACCCCCACGACCCCCCACAA 3850 40 T A T H P D D P D D P T N P H N CACACCCACACGCACCCACACACAAACCACACGCGTCCTCACCGCCCTCC 3900 TPTRTHTQTTRULTAL AACACCACCTCATCACCACCAACCACACCCTCATCGTCCACACCACCACCACC 3950 45 Q H H L I T T N H T L I V H T T T DPPGAAVTGLTRTAQNE ACACCCGGCCGCATCCACCTCATCGAAACCCACCACCCCCACACCCCCAC 4050 HPGRIHLIETHHPHTP 50 TCCCCCTCACCCAACTCACCACCTCCACCAACCCCACCTACGCCTCACC 4100 L P L T O L T T L H O P H L R L T NNTLHTPHLTPITTHHN 55 T T T T P N T P P L N P N H A ILITGGSGTLAGILARH CTCAACCACCCCACACCTACCTCTCTCCCGCACACCACCACCCCCCAC 4300 LNHPHTYLLSRTPPPT 60 T P G T H I P C D L T D P T Q I CCCAAGCCCTCACCCACATACCACAACCCCTCACCGGCATCTTCCACACC 4400 TQALTHIPOPLTGIFHT GCCGCCACCCTCGACGACGCCACCCTCACCAACCTCACCCCCCCAACACCT 4450

A A T L D D A T L T N L T P Q H L CACCACCACCCTCCAACCCAAAGCCGACGCCTGGCACCTCCACCACC 4500 T T T L Q P K A D A A W H L H H ACACCCAAAACCAACCCCTCACCCACTTCGTCCTCTACTCCAGCGCCGCC 4550 H T O N O P L T H F V L Y S S A A GCCACCCTGGGCAGCCCCGGGCAAGCCAACTACGCCCAACGCCTT 4600 ATLGSPSQANYAAANAF COTEGACGCOCTOGCCACCCACCCCACACCCAAGGACAACCCGCCACCA 4600 LDALATHRHTQSQPAC 10 DENTEGESTOGGGENTGTGGCACACCACCACCACACACTCHCCAGCCAACTC 4700 TIAWGMWHTTTTLTSQL ACCCACACOGOGOGOGOGOGOGGGGGCTTCCTGCCGATCTC 4750 T D S D R D R I R R G G F L P I S GGACGACGAGGGCATGC 15 D D E G M

The *Nhel-Xhol* hybrid FK-506 PKS module 8 containing the AT domain of module 12 of rapanyein is shown below.

GCATGCGGCTGTACGAGGCGGCACGGCACCGGAAGTCCCGTGGTGGTG 50 M R L Y E A A R R T G S P V V V 20 GOGGCCGCTCGACGACGCCCCGGACGTGCCGCTGCTGCGCGGGCTGCG 100 A A A L D D A P D V P L L R G L R GCGTACGACCGTCCGGCGTGCCGCCGTCCGGGAACGGTCTCTCGCCGACC 180 R T T V R R A A V R E R S L A D 25 GCTCGCCGTGCTGCCGACGACGASGGCGCGACGCCTCCCTCGCGTTCG 200 RSPOOPTTSAPTPPSRS TOOTGGAACAGCGCGACGTGGTCGGCCACCTGGGCGCCGAAGACAT 250 S W N S T A T V L G H L G A E D I CCCGGCGACGACGACGTTCAAGGAACTCGGCATCGACTCGCTCACCGCGG 300 30 PATTFKELGIDSLTA TCCAGCTGCGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGCC 350 V Q L P N A L T T A T G V R L N A T A V F D F P T P R A L A A R L G 35 CGACGASCTGGCCGGTACCCGCGCGCCCGTCGCGGCCCGGACCGCGCCCA 450 DELAGTRAPVAARTAA CCGCGGCGCGCACGACGAACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500 T A A A H D E P L A I V G M A C R CTGCCGGGCGGGTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550 40 L P G G V A S P Q E L W R L V A S CGGCACGGACGCCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG 600 G T D A I T E F P A D R G W D V ACGCGCTCTACGACCCGGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650 DALYDPDPDAIGKTFVR 45 CACGGCGGCTTCCTCGACGGTGCGACCGGCCTTCGACGCGGCGTTCTTCGG 700 H G G F L D G A T G F D A A F F G GATCAGCCCGCGGAGGCCCTGGCCATGGACCCGCAGCAACGGGTGCTCC 750 I S P R E A L A M D F Q Q R V L TGGAGACGTCCTGGGAGGCCTTCGAAAGCGCGGGCATCACCCCGGACGCG 800 50 L E T S W E A F E S A G I T P D A GCGCGGGCAGCGACACCGGCGTGTTCATCGGCGCGTTCTCCTACGGGTA 850 ARGSDTGVFIGAFSYGY CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTCGCAGACCA 900 G T G A D T N G F G A T G S Q T GCGTGCTCTCCGGGCCGCCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950 55 S V L S G R L S Y F Y G L E G P S GTCACGGTCGACACCGCCTGCTCGTCGTCACTGGTCGCCCTGCACCAGGC 1000 V T V D T A C S S S L V A L H Q A AGGGCAGTCCCTGCGCTCGGCGAATGCTCGCTCGCCCTGGTCGGCGGTG 1050 60 G Q S L R S G E C S L A L V G G TCACGGTGATGGCGTCGCCGGGGGGATTCGTCGAGTTCTCCCGGCAGCGC 1100 V T V M A S P G G F V E F S R Q R

	GGGCTGGCGGCGGCGGGGGGGGGGGGGGGGGGGGGGGG	1150
	TACGAGETTEGGEGAGGGGGGGGGGGGGGGGGGGGGGGGG	1200
5	ACGOSGAGOGGAGAGOGTCCTCSCCCTCGTACGCGGGTCCGGGCCTCCGGGCCTCCGGGCCTCGGGGGGGG	
	ANSDOASNGLSAFNGPS	1300
10	QERVIHQALANAKLTP	1350
	CODATOTOGACGOSTOGAGGCGCACGGCACCGGGCACCGGCGCGCGACACCGGGCACCGGCACCAC	1400
	CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC F I E A Q A L L A T Y G Q D R A T	1450
15	GCCCCTGCTGCTGGCTGGCTGAGGTCGAACATCGGGGGGGCGAGGCCG P L L L G S L K S N I G H A Q A	1500
	CGTCAGGGGTCGCCGGGATCATCAAGAMGGTGCAGGCCATCCGGCACGGG	1550
20	GAACTGCCGCGACACGCGGACGACGCGCCGCCGCCGCCGCCGCCG	1600
	GACGGCCGGTGGAGCTCCTGACGTCGGCCGGGGGATCGACGGGGATCGACGTCGACGTCGGCGGGGATCGACGTCGGCCGGGGGATCGACGTCGGCCGGGGGATCGACGTCGGCCGGGGGATCGACGTCGACGTCGGCCGGGGGATCGACGTCGGCCGGGGGATCGACGTCGGCCGGGGGATCGACGTCGACGTCGACGTCGGCCGGGGGATCGACGTCGACGTCGACGTCGACGTCGACGTCGGCCGGGGGATCGACGGGGGATCGACGTCACACACA	1650
	COGGTCGCCCGCGCGCGCGCTGCCGTCTCGTCGTTCGCCGTGAGCGGCACG T G R P R R A A V S S F G V S G T	1700
25	AACGCCCACATCATCCTTGAGGCAGGACCGGTCGA N A H I I L E A G P V K T G P V E	1750
	GGCAGGAGCGATCGAGGCAGGACCGGTCGAGGAGCAGGCTG A G A I E A G P V E V G P V E A	1800
30		1850
30	CTCGTGTCGGCGCGTTCCCCGGGAGGCACTCGACGAGCAGATCGGGCGCCT	1900
	L V S A R S P E A L D E Q I G R L GCGCGCCTATCTCGACACCGGCCCGGGCGTCGACCGGGCGCGTGGCGCGCGC	1950
35	AGACACTGGCCGGGCTACGCACCTGCCGGGCCGTACTGCTCGGG	2000
	Q T L A R R T H F T H R A V L L G	
	GACACCGTCATCGGCGCTCCCCCCGCGGACCAGGCCGACGAACTCGTCTT D T V I G A P P A L Q A D E L V F	
40	CGTCTACTCCGGTCAGGCACCCAGCATCCCGGGATGGGGGAGCAGCTAG V Y S G Q G T Q H P A M G E Q L	2100
	CCGCCGCGTCTTCGCGCGGATCCATCAGCAGGTGTGGGACCTG A A A F P V F A R I H Q Q V W D L	2150
	CTCGATGTGCCCGATCTGGAGGTGAACGAGACCGGTTACGCCCAGCCGGC L D V P D L E V N E T G Y A Q P A	2200
45	CCTGTTCGCAATGCAGGTGGCTCTGTTCGGGCTGCTGGAATCGTGGGGTG	2250
	TACGACCGGACGCGGTGATCGGCCATTCGGTGAGCTTGCGGCTGCG	2300
5.0	TATGTGTCCGGGCTTGTGGTCGTTGGAGGATGCCTGCACTTTGGTGTCGCC	2350
50	Y V S G V W S L E D A C T L V S A GOGGGCTCGTCTGATGCAGGCTCTGCCCGCGGGGTGGGGT	2400
	R A R L M Q A L P A G G V M V A TCCCGGTCTCGGAGGATGAGGCCCGGGCCGTGCTGGGTGAGGGTGTGGAG	2450
55	V P V S E D E A R A V L G E G V E ATCGCCGCGGTCAACGGCCCGTCGTCGGTGGTTCTCTCCGGTGATGAGGC	2500
	I A A V N G P S S V V L S G D E A CGCCGTGCTGCAGGCCGCGGAGGGGCTGGGGAAGTGGACGCGGCTGGCGA	2550
	A V L Q A A E G L G K W T R L A CCAGCCACGCGTTCCATTCCGCCCGTATGGAACCCATGCTGGAGGAGTTC	2600
60	T S H A F H S A R M E P M L E E F CGGGCGGTCGCCGAAGGCCTGACCTACCGGACGCCGCAGGTCTCCATGGC	
	R A V A E G L T Y R T P Q V S M A CGTTGGTGATCAGGTGACCACCGCTGAGTACTGGGTGCGGCAGGTCCGGG	2700
	V G D Q V T T A E Y W V R Q V R	

	ACACGSTCCGGTTCGGCGAGCAGGTGGCCTCGTACGAGGACGCCGTGTTC D T V R F G E Q V A S Y E D A V F	2750
	GTCGAGCTGGGTGGCGACCGGTCACTGGCCCGCCTGGTCGACGGTGTCGC V S L G A D R S L A R L V D G V A	2800
5	GATGCTGCACGGCGACCACGAAATCCAGGCCGCGATCGGCCGCCCCCCCC	2850
	ACCTGTATGTCAACGGCGTCACGGTCGACTGGCCGCGCGCTCCTGGGCGAT H L Y V N G V T V D W P A L L G D	2900
10	GCTCCGGCAACACGGGTGCTGGACCTTCCGACATACGCCTTCCAGCACCA A P A T R V L D L P T Y A F Q H Q	2950
	GOGOTACTGGCTCGGGCTCCCCCGGGCCACGGCCGACTCGGGCCACC R Y W L E S A P P A T A D S G H	3000
	CCGTCCTCGGCACTCGCCGTCGCCGGGTCGCCGGCCGGGTGTTC P V L G T G V A V A G S F G R V F	3050
15	ACGGGTCCCGTGCCGCGGGTGCGGACCGCGGGGGTGTTCATCGCCGAACT T G P V P A G A D R A V F I A E L	3100
		3150
20	ACGTCACCTCCGTGCCGGGGGGATCCGCCGGGGCACGGCACGACG	3200
	ACCTGGGTCGATGAACCGGCGGCGGCGGCGGCGGCGCGCTCACCGTCCA T W V D E P A A D G R R R F T V H	3250
	CACCCGCGTCGGCCGCGCGCGGGCGGGGGGGTTCTCC T R V G D A P W T L H A E G V L	3300
25	GCCCCGGCCGCGTGCCCCAGCCGGAAGCCGTCGACACGGCCTGGCCCCCG R P G R V P Q P E A V D T A W P P	3350
		3400
30	CCAGGTCTTCGTCGAAGCCGAAGTCGACAGCCCTGACGCCTTCGTGGCAC	3450
	ACCCCGACCTGCTCGACGCGGTCTTCTCCGCGGTCGGCGACGGGACGGGACGGAC	3500
	CAGCCGACCGGATGGCGCGACCTCGCGGTGCACGCGTCGGACGCCACCGT Q P T G W R D L A V H A S D A T V	3550
35	GCTGCGCGCCTGCCTCACCCGCCGCGACAGTGGTGTCGTGGAGCTCGCCG L R A C L T R R D S G V V E L A	3600
	CCTTCGACGGTGCCGGAATGCCGGTGCTCACCGCGGAGTCGGTGACGCTG A F D G A G M P V L T A E S V T L	3650
40		3700
	GOTTGAGTGGTTGCCGGTGGCGGAGGCCCACTACGACGGTGCCGACGAGG L E W L P V A E A H Y D G A D E	3750
	TGCCCGAGGGCTACACCCTCATCACCGCCACACACCCCGACGACGCCGACGACCCCGAC	3800
45	GACCCCACCAACCCCACACACCCACACGCACCCACACACAAACCAC D P T N P H N T P T R T H T Q T T	3850
	ACGCGTCCTCACCGCCCTCCAACACCCACCTCATCACCACCAACCA	3900
50	TCATCGTCCACACCACCACCGACCCCCCAGGCGCCGCCGTCACCGGCCTC L I V H T T T D P P G A A V T G L	3950
	ACCCGCACCGCACAAAACGAACACCCCGGCCGCATCCACCTCATCGAAAC T R T A Q N E H P G R I H L I E T	
	CCACCACCCCACACCCCACTCCCCCTCACCCAACTCACCAC	4050
55	AACCCCACCTACGCCTCACCAACAACACCCTCCACACCCCCACCTCACC	4100
	CCCATCACCACCACCACACACACCACCACCACCACCCCCAACACCCC	
60	CCTCAACCCCAACCACGCCATCCTCATCACCGGGGGGTCCGGGCACCCTCG	4200
	CCGGCATCCTCGCCCGCCTCCTCAACCACCCCCACACCTACCT	4250
	CGCACACCACCACCACACACCCGGCACCCACATCCCCTGCGACCT R T P P P P T T P G T H I P C D L	4300

CACCGACCCCACCCAAATCACCCAAGCCCTCACCCACATACCACAACCCC 4350 TOPTQITQALTHIPQP TOACUGGOATOTTOGACACCGGGGAACCTCGACGACGACGCCACCCTCACC 4400 GIFHTAATLUDATLT ANCOTORCOSSCARGACOTORSCASCASCTSCARGOSFARGOSGAGGO 4450 N L T P Q H L T T T L D P H A D A GGCCTGGCACCTCCACCACCACACCCAAAACCAACCCCTCACCCACTTCG 4503 AWHLHHHTONQPLTHF TCCTCTACTCCAGCGCGGCGGCCACCCTCGGCAGCCCGGGCCAAGCCAAC 4550 10 V L Y S S A A A T L G S P G Q A N Y A A A N A F L D A L A T H R H T CCAAGGACAACCGCCACCACCATCGCCTGGGGCATGTGGGCACCACCACCA 4650 Q G Q P A T T I A W G M W H T T 15 COACACTCACCAGCCAACTCACCGACAGCGACCGGGACCGCATCCGCCGC 4700 T T L T S Q L T D S D R D R I R R GGCGGCTTCCTGCCGATCTCGGACGACGAGGGCATGC G G F L P I S L D E G M

The *NheI-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module 13 of rapamycin is shown below.

GCATGCGGCTGTACGAGGCGGCACGGCACCGGAAGTCCCGTGGTGGTG 50 M R L Y E A A R R T G S P V V V GCGGCCGCTCGACGACGCCCCCGGACGTGCCGCTGCT%U4CGGGCTGCG 100 25 A A A L D D A P D V P L L R G L R GOSTACGACCGTCCGGCGTGCCGCCGTCCGGGAACGCTCTCTCGCCGACC 150 R T T V R R A A V R E R S L A D R S P C C P T T S A P T P P S R S TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250 30 SWNSTATVLGHLGAEDI CCCGGCGACGACGTTCAAGGAACTCGGCATCGACTCGCTCACCGCGG 300 PATTFKELGIDSLTA TCCAGCTGCGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGCC 350 35 V Q L R N A L T T A T G V R L N A TAVFDFPTPRALAARLG CGACGAGCTGGCCGGTACCCGCGCGCCCGTCGCGGCCCGGACCGCGGCCCA 450 D E L A G T R A P V A A R T A A 40 CCGCGGCGCGCACGACGACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500 T A A A H D E P L A I V G M A C R CTGCCGGGCGGGGTCGCCTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550 L P G G V A S P Q E L W R L V A S CGGCACCGACGCCATCACGGAGTTCCCCGCGGGCCGCCTGGG7 naTGG 600 45 G T D A I T E F P A D R G W D V ACGCGCTCTACGACCCGGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650 D A L Y D P D P D A I G K T F V R CACGGGGGCTTCCTCGACGGTGCGACGGGCTTCGACGCGGGGTTCTTCGG 700 H G G F L D G A T G F D A A F F G 50 GATCAGCCGCGCGAGGCCCTGGCCATGGACCCGCAGCAACGGGTGCTCC 750 I S P R E A L A M D P Q Q R V L TGGAGACGTCCTGGGAGGCGTTCGAARGCGCGGCATCACCCCGGACGCG 800 LETSWEAFESAGITPDA GCGCGGGGCAGCGACACCGGCGTGTTCATCGGCGCGTTCTCCTACGGGTA 850 55 ARGSDTGVFIGAFSYGY CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTCGCAGACCA 900 G T G A D T N G F G A T G S Q T GCGTGCTCTCCGGCCGCCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950 S V L S G R L S Y F Y G L E G P S 60 GTCACGGTCGACACCGCCTGCTCGTCGTCACTGGTCGCCCTGCACCAGGC 1000 V T V D T A C S S S L V A L H Q A AGGGCAGTCCCTGCGCTCGGCGAATGCTCGCTCGCCCTGGTCGGCGGTG 1050

G Q S L R S G E C S L A L V G G TCACGGTGATGGCGTCGCCCGGCGGGATTCGTCGAGTTCTCCCGGCAGCGC 1100 V.T V M A S P G G F V E F S R Q R 5 G L A P D G R A K A F G A G A D G T S F A E G A G A L V V E R L S AGGGGGAGGGCCACGGCCACACGGTCCTCGCCCTCGTACGGGGGCTCCGCG 1250 DAERHGHTVLALVRGSA 10 GCTAACTCCGACGCGCGTCGAACGGTCTGTCGGCGCCGAACGGCCCCTC 1300 A N S D G A S N G L S A P N G P S CCAGGAACGCGTCATCCACCAGGCCCTCGCGAACGCGAAACTCACCCCCG 1350 Q E R V I H Q A L A N A K L T P 15 A D V D A V E A H G T G T R L G D CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450 PIEAQALLATYGQDRAT GOCCCTGCTGCTCGCCTCACTGAAGTUCATCGGGCACGCCCAGGCCG 1500 PL L G S L K S N I G H A Q A 20 CGTCAGGGGCCGGGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550 ASGVAGIIKMVQAIRHG GAACTGCCGCCGACACTGCACGCGGGACGGTCGCCGCACGTCGACTG 1600 E L P P T L H A D E P S P H V D W 25 TAGAVELLTSARPWPG CCGGTCGCCGCGCGCGCGCTGCCGTTCGTCGTTCGGCGTGAGCGGCACG 1700 T G R P R R A A V S S F G V S G T AACGCCCACATCATCCTTGAGGCAGGACCGGTCAAAACGGGACCGGTCGA 1750 NAHIILEAGPVKTGPVE 30 GGCAGGAGCGATCGAGGCAGGACCGGTCGAAGTAGGACCGGTCGAGGCTG 1800 A G A I E A G P V E V G P V E A GACCGCTCCCCGCGGCGCCGCCGTCAGCACCGGGGGAAGACCTTCCGCTG 1850 G P L P A A P P S A P G E D L P L CTCGTGTCGGCGCGTTCCCCGGAGGCACTCGACGAGCAGATCGGGCGCCT 1900 35 LVSARSPEALDEQIGRL GCGCGCCTATCTCGACACCGGCCCGGGCGTCGACCGGGCGGTCGTGGCGC 1950 RAYLD T G P G V D R A A V A AGACACTGGCCCGGCGTACGCACTTCACCCACCGGGCCGTACTGCTCGGG 2000 Q T L A R R T H F T H R A V L L G 40 GACACCGTCATCGGCGCTCCCCCCCGGGGCCAGGCCGACGAACTCGTCTT 2050 D T V I G A P P A D Q A D E L V F CGTCTACTCCGGTCAGGGCACCCAGCATCCCGCGATGGGCGAGCAGCTAG 2100 V Y S G Q G T Q H P A M G E Q L CCGATTCGTCGGTGTTTCGCCGAGCGGATGGCCGAGTGTGCGGCGCGCG 2150 45 A D S S V V F A E R M A E C A A A TTGCGCGAGTTCGTGGACTGGGATCTGTTCACGGTTCTGGATGATCCGGC 2200 LREFVDWDLFTVLDDPA GGTGGTGGACCGGGTTGATGTGGTCCAGCCCGCTTCCTGGGCGATGATGG 2250 V V D R V D V V Q P A S W A M M 50 TTTCCCTGGCCGCGGTGTGGCAGGCGGCCGGTGTGCGGCCGGATGCGGTG 2300 V S L A A V W Q A A G V R P D A V ATCGGCCATTCGCAGGGTGAGATCGCCGCAGCTTGTGTGGCGGGTGCGGT 2350 I G H S Q G E I A A A C V A G A V GTCACTACGCGATGCCGCCCGGATCGTGACCTTGCGCAGCCAGGCGATCG 2400 55 SLRDAARIVTLRSQAI CCCGGGGCCTGGCGGGCCGGGGGGGGCGATGGCATCCGTCGCCCTGCCCGCG 2450 A R G L A G R G A M A S V A L P A CAGGATGTCGAGCTGGTCGACGGGGCCTGGATCGCCGCCCACAACGGGCC 2500 Q D V E L V D G A W I A A H N G P 60 CGCCTCCACCGTGATCGCGGGCACCCCGGAAGCGGTCGACCATGTCCTCA 2550 A S T V I A G T P E A V D H V L CCGCTCATGAGGCACAAGGGGTGCGGGTGCGGCGGATCACCGTCGACTAT 2600 T A H E A Q G V R V R R I T V D Y GCCTCGCACACCCCGCACGTCGAGCTGATCCGCGACGAACTACTCGACAT 2650

	A S H T P H V E L I R D E L L D I CASTAGOGAGAGGAGGAGGAGGAGGAGGAGGGGGGGGGGGG	2700
	T S D S S S Q T P L V P W L S T	2700
	TagAdagcAdmaguragAdAgcddgcTagAdaggAgTAdTagTACCag	2750
5	TO SOME WORK DOES FOR DOGS EVEN YORK	2,50
-	AACOMOCOTOAACCGGTTTCCACCCCCCCCTCAGCTAGTTGCAGGC	1800
	N L R E P V G F H P A V S Q L Q A	
		2850
	Q G D T V F V E V S A S P V L L	
10	AGGCGATGGACGACGATGTCGTCACGGTTGCCACGCTGCGTGGTGACGAC	2900
	Q A M D D D V V T V A T L R R D D	
	GGCCACCCCGATGCTCACCGCCCTGGCACAGGCCTATGTCCACGG	2950
	G D A T R M L T A L A Q A Y V H G	
	CGTCACCGTCGACTGGCCCGCCATCCTCGGCACCACACAACCCGGGTAC	3000
15	V T V D W P A I L G T T T R V	
	TUGACCTTCCGACCTACGCCTTCCAACACCAGCGGTACTGGCTCGAGTCG	3050
	L O L P T Y A F Q H Q R T T S	
	GOT COCCOGGCOACGGCCACTCGGGCCACCCGTCU: CGGUALL	3107
20	A P P A T A D S G H P V L G T G V	5560
20	SCOOTSCOGGSCOGGSCOGGSTCTTCACGGGTCCCGTGCCCGCCG A V A G S P G R V F T G P V P A	3150
		3200
	G A D R A V F I A E L A L A A A D	0200
	GOCACCBACTOCBCCACGGTCGAACAGCTCGACGTCACCTCCGTGCCCGG	3250
25	A T D C A T V E Q L D V T S V P G	
		3300
	3 S A R G R A T A Q T W V D E P	
	CCCCCGACGGGGGGGCGCTTCACCGTCCACACCCGGGTCGGCGACGCC	3350
	A A D G R R R F T V H T R V G D A	
30	0001360301000000366001.	3400
	PWTLHAEGVLRPGRVPQ	2450
	GCCCGAAGCCGTCGACACCGCCTGGCCCCCGCCGGGCGGG	3450
	PEAVDTAWPPPGAVPA ACGSSTGCCCGGGGGCTGGCGAGGCGAGGTCTTCGTCGAAGCC	3500
35	D G L P G A W R R A D O V F V E A	5500
33		3550
	EVDSPDGFVAHPDLLDA	
		3600
	V F S A V G D G S R Q P T G W R	
40	ACCTCGCGGTGCACGCGTCGGACGCCACCGTGCTGCGCGCGC	3650
	D L A V H A S D A T V L R A C L T	
	CGCCGCGACAGTGGTGGTGGAGCTCGCCGCCTTCGACGGTGCCGGAAT	3700
	R R D S G V V E L A A F D G A G M	225
4.5	GCCGGTGCTCACCGCGGAGTCGGTGACGCTGGGCGAGGTCGCCAG	3/50
45	PVLTAESVTLGEVASA	3200
	GCGGATCCGACGACTCGGACGGTCTGCTTCGCCTTGAGTGGTTGCCGGTG	3500
	G G S D E S D G L L R L E W L P V GCGGAGGCCCACTACGACGGTGCCGACGAGGTGCCCGAGGGCTACACCCT	3850
	A E A H Y D G A D E L P E G Y T L	2030
50	CATCACCGCCACACCCCGACGACGCCGACGACCCCACCAACCCCCACA	3900
50	I T A T H P D D P D D P T N P H	
	ACACACCCACACCCACACACACACACACACACGCGTCCTCACCGCCCTC	3950
	NTPTRTHTQTTRVLTAL	
	CAACACCACCTCATCACCACCAACCACCCTCATCGTCCACACCACCACCAC	4000
55	Q H H L I T T N H T L I V H T T T	
	CGACCCCCAGGCGCCGTCACCGGCCTCACCCGCACCGCA	4050
	D P P G A A V T G L T R T A Q N	4100
	AACACCCCGGCCGCATCCACCTCATCGAAACCCACCCCCCACACCCCCA	4100
60	E H P G R I H L I E T H H P H T P	4150
60	CTCCCCCTCACCCAACTCACCACCCTCCACCAACCCACCTACGCCTCAC	510U
	CAACAACACCCTCCACCCCCCCCCCCCCCCATCACCACCCCCACCA	4200
	N N T L H T P H L T P I T T H H	3200
	ACACCAGCACAACCACCCCAACACCCCCACCCCTCAACCCCAACCACGCC	4250
	MCACCACCACAACCACCCCCACACACCCCCCCCCCCCC	

NTTTTPNTPPLNPNHA I-L I T G G S G T L A G I L A R H OT CARCOROGOGA CACCTACCTCCTCCCCCCAAACCACCACCCCCA 4350 5 L M H P H T Y L L S R T P P P P COMBRIGOCOGOR DOCACATO DOCATOGORA DOTAR CORROCORROCORRANTO E 44 00 T T P G T H I P G S L T D P T Q ACCORAGODOTOROSOROATROCAGARGOSOTOROGGGATOTTOCAGRO 4450 T Q A L T H I F Q P L T G I F H T 10 TGGCGCCACCCTCGACGACGCCACCTCACCAACCCCCCAACACC 4500 A A T L D D A T L T N L T P Q H TCACCACCACCCTCCAACCCAMAGOOGACGCCGCCTGGCACCTCCACCAC 4550 LTTTLQPKADAAWHLHH CROAD CARAROCARCOCTOROCTACTTOGTCCTCTACTCCAGCGCCGC 4600 H T Q N Q P L T H F V L Y S S A A 15 0300A000T0000Ag0000Gg00AAG00AACTA000G00000AACG00T 4650 ATLGSPGQANYAAANA Toomigacgccctcgccannnaccg:accolargacaacccgccacc 4000 F L D A L A T H A H T Q G Q P A T 20 ACCATOGOCTGGGGCATGTGGCACACCACCACCACTCACCAGCCAACT 4750 TIAWGMWHTTTTLTSQL CACCGACAGCGACCGCATCCGCCGCGCGCGCGCGCTTCCTGCCGATCT 4800 T D S D R D R I R R G G F L P I DGGACGACGAGGGCATGC 25 SDDEGM

30

35

40

45

Example 3

Recombinant PKS Genes for 13-desmethoxy FK-506 and FK-520

The present invention provides a variety of recombinant PKS genes in addition to those described in Examples 1 and 2 for producing 13-desmethoxy FK-506 and FK-520 compounds. This Example provides the construction protocols for recombinant FK-520 and FK-506 (from *Streptomyces* sp. MA6858 (ATCC 55098), described in U.S. Patcnt Nos. 5,116,756, incorporated herein by reference) PKS genes in which the module 8 AT coding sequences have been replaced by either the *rap*AT3 (the AT domain from module 3 of the rapamycin PKS), *rap*AT12, *ery*AT1 (the AT domain from module 1 of the erythromycin (DEBS) PKS), or *ery*AT2 coding sequences. Each of these constructs provides a PKS that produces the 13-desmethoxy-13-methyl derivative, except for the rapAT12 replacement, which provides the 13-desmethoxy derivative, i.e., it has a hydrogen where the other derivatives have methyl.

Figure 7 shows the process used to generate the AT replacement constructs. First, a fragment of ~4.5 kb containing module 8 coding sequences from the FK-520 cluster of ATCC 14891 was cloned using the convenient restriction sites SacI and SphI (Step A in Figure 7). The choice of restriction sites used to clone a 4.0 - 4.5 kb fragment comprising module 8 coding sequences from other FK-520 or FK-506 clusters can be different depending on the DNA sequence, but the overall scheme is identical. The unique SacI and SphI restriction sites at the ends of the FK-520 module 8 fragment were then changed to unique Bgl II and NsiI sites by ligation to synthetic linkers (described in

the preceding Examples, see Step B of Figure 7). Fragments containing sequences 5' and 3' of the AT8 sequences were then amplified using primers, described above, that introduced either an *Avr*II site or an *Nhe*I site at two different KS'AT boundaries and an *Nho*I site at the AT DH boundary (Step C of Figure 7). Heterologous AT domains from the rapamycin and erythromycin gene clusters were amplified using primers, as described above, that introduced the same sites as just described (Step D of Figure 7). The fragments were ligated to give hybrid modules with in-frame fusions at the KS/AT and AT/DH boundaries (Step E of Figure 7). Finally, these hybrid modules were ligated into the *Bam*HI and *Pst*I sites of the KC515 vector. The resulting recombinant phage were used to transform the FK-506 and FK-520 producer strains to yield the desired recombinant cells, as described in the preceding Examples.

5

10

15

The following table shows the location and sequences surrounding the engineered site of each of the heterologous AT domains employed. The FK-506 hybrid construct was used as a control for the FK-520 recombinant cells produced, and a similar FK-520 hybrid construct was used as a control for the FK-506 recombinant cells.

Heterologous AT	Enzyme	Location of Engineered Site
FK-506 AT8	AvrII	GGCCGT <u>bagba</u> cCGTCCGGCGGTCTCGTCGTTC
(hydroxymalonyl)		GRPRRAAVSSF
, ,	NheI	ACCCAGCATCCCSCGATGGGTGAGCG <u>actagc</u> C
	1 11/161	TQHPANGERLA
		TACGCCTTCCAGCGGCGGCCCTACTGGatcgag
	Ahol	YAFÇRRPYWIE
rapamycin AT3	AvrII	GACCGG <u>cuccat</u> CGGGCGGGGGGGTGTCGTTC
(methylmalonyl)	i	DRPPRAGVSSF
-	Nhel	TGGCAGTG3CTGGGGAT3CGCASTGCcctgcgG
	,	WQWLGMGSALR
	XhoI	TACGCCTTCCAACACCAGCGGTACTGGgtegag
		Y A F Q H Q R Y W V E
rapamycin AT12	AvrII	GGCCGAdagaaaCGGGCAGGCGTGTCGTCCTTC
(malonyl)		G R A R R A G V S S F
	NheI	TCGCAGCGTGCTGGCATGGGTGAGGAactggcC
	İ	S Q K A G M G E E L A
	XhoI	TACGCCTTCCAGCACCAGCGCTACTGGctogag
DEDC ATI		Y A F Q H Q R Y W L E
DEBS AT1	AvrII	GCGCGAddadadcGGGCGGGGGGTCTCGTCGTTC
(methylmalonyl)		A R P R R A G V S S F
	Nhel	TGGCAGTGGGCGGCATGGCCGTCGAcctgctC
1		W Q W A G M A V D L L TACCOGTTCCAGCGCGAGCGCGTCTGGctcgaa
	XhoI	
DEBS AT2		Y F F Q R E R V W L E GACGGGgtgggCAGGTGTGTCGGCGTTC
	AvrII	D G V R R A G V S A F
(methylmalonyl)		GCCCAGTGGSAAGGCATGGCGGGGAGttgttG
	_	GCCCAG 1 GGCAAAGCA 1 GGCGCGGAAA C CA CCG

Nh	el A	Q	W	E	G	М	А	R	E	L	L
	TATCCTTTCCAGGGCAAGCGGTTCTGGetgetg										
Xh	oI Y	P	F	Q	G	K	R	F	W	L	L

The sequences shown below provide the location of the KS/AT boundaries chosen in the FK-520 module 8 coding sequences. Regions where *AvrII* and *NheI* sites were engineered are indicated by lower case and underlining.

5 A S A V E L L T S A B F W P E T D R P R GTGCCGCCCTCCTCCTCCTTCGCCCCTGAGCGGCACCAACGCCCAACGCCCACGTCATCCTGGAGGCCG R A A V S S F G V S G T N A H V I L E A GACCSSTARCGSAGACGCCCGCGGCATCGCCTTCCGGTGACCTTCCCCCTGCTGGTGTCGG G P V T E T P A A S P S G D L P L L V S 10 CACGCTCACCGGAAGCGCTCGACGAGCAGATCCGCCGACTGCGCGCCTACCTGGACACCA A R S P E A L D E Q I R R L R A Y L D T CCCCGGACGTCGACCGGGTGGCCGTGGCACAGACCCTGGCCCGGCGCACACACTTCGCCC T P D V D R V A V A Q T L A R R T H F A 15 H R A V L L G D T V I T T P P A D R P D ELVFVYSGQGTQHPAMGEQL cCGCCGCCCATCCCGTGTTCGCCGACGCCTGGCATGAAGCGCTCCGCCGCCTTGACAACC A A A P P V F A D A W H E A L R R L D N 20

The sequences shown below provide the location of the AT/DH boundary chosen in the FK-520 module 8 coding sequences. The region where an *XhoI* site was engineered is indicated by lower case and underlining.

TCCTCGGGGCTGGGTCACGGCACGACGCGGATGTGCCCGCGTACGGGTTCCAACGGCGGC

I L G A G S R H D A D V P A Y A F Q R R ACTACTGGatcgagTCGGCACGCCGGCCGCGCACCGGGGGCCACCCCGTGCTGGGCT H Y W I E S A R F A A S D A G H P V L G

30

The sequences shown below provide the location of the KS/AT boundaries chosen in the FK-506 module 8 coding sequences. Regions where *Avr*II and *Nhe*I sites were engineered are indicated by lower case and underlining.

TCGGCCAGGCCGTGGCCGGACCGGCCGTgcggcgCGTGCGGCGGTCTCGTCGTTCGGG SARPWPRTGRPRRAAVSSFG 35 V S G T N A H I I L E A G P D Q E E P S GCAGAACCGGCGGTGACCTCCCGCTGCTGTGTCGGCACGGTCCCCGGAGGCACTGGAC A E P A G D L P L L V S A R S P E A L D GAGCAGATCGGGCGCTGCGCGACTATCTCGALGCCCCCCGGCGTGGACCTGGCGCGCC EQIGRLRDYLDAAPGVDLAA 40 GTGGCGGGGACACTGGCCACGCGTACGCACTTCTCCCACCGCGCGTACTGCTCGGTGAC V A R T L A T R T H F S H R A V L L G D ACCGTCATCACCGCTCCCCCCGTGGAACAGCCGGGCGAGCTCGTCTTCGTCTACTCGGGA T V I T A P P V E Q P G E L V F V Y S G CAGGGCACCCAGCATCCCGCGATGGGTGAGCGgetcgcCGCAGCCTTCCCCGTGTTCGCC 45 QGTOHPAMGERLAAAFPVFA GACCCGGACGTACCCGCCTACGCCTTCCAGCGGCCGCCCTACTGGATCGAGTCCGCGCCCG D P D V P A Y A F Q R R P Y W I E S A P

The sequences shown below provide the location of the AT/DH boundary chosen in the FK-506 module 8 coding sequences. The region where an *XhoI* site was engineered is indicated by lower case and underlining.

Example 4

Replacement of Methoxyl with Hydrogen or Methyl at C-15 of FK-506 and FK-520

The methods and reagents of the present invention also provide novel FK-506

and FK-520 derivatives in which the methoxy group at C-15 is replaced by a hydrogen or methyl. These derivatives are produced in recombinant host cells of the invention that express recombinant PKS enzymes the produce the derivatives. These recombinant PKS enzymes are prepared in accordance with the methodology of Examples 1 and 2, with the exception that AT domain of module 7, instead of module 8, is replaced. Moreover, the present invention provides recombinant PKS enzymes in which the AT domains of both modules 7 and 8 have been changed. The table below summarizes if e various compounds provided by the present invention.

	Compound	C-13	C-15	Derivative Provided
15	FK-506	hydrogen	hydrogen	13, 15-didesmethoxy-FK-506
	FK-506	hydrogen	methoxy	13-desmethoxy-FK-506
	FK-506	hydrogen	methyl	13,15-didesmethoxy-15-methyl-FK-506
	FK-506	methoxy	hydrogen	15-desmethoxy-FK-506
	FK-506	methoxy	methoxy	Original Compound FK-506
20	FK-506	methoxy	methyl	15-desmethoxy-15-methyl-FK-506
	FK-506	methyl	hydrogen	13,15-didesmethoxy-13-methyl-FK-506
	FK-506	methyl	methoxy	13-desmethoxy-13-methyl-FK-506
	FK-506	methyl	methyl	13,15-didesmethoxy-13,15-dimethyl-FK-506
	FK-520	hydrogen	hydrogen	13, 15-didesmethoxy FK-520
25	FK-520	hydrogen	methoxy	13-desmethoxy FK-520
	FK-520	hydrogen	methyl	13,15-didesmethoxy-15-methyl-FK-520
	FK-520	methoxy	hydrogen	15-desmethoxy-FK-520
	FK-520	methoxy	methoxy	Original Compound FK-520
	FK-520	methoxy	methyl	15-desmethoxy-15-methyl-FK-520
30	FK-520	methyl	hydrogen	13,15-didesmethoxy-13-methyl-FK-520
	FK-520	methyl	methoxy	13-desmethoxy-13-methyl-FK-520
	FK-520	methyl	methyl	13,15-didesmethoxy-13,15-dimethyl-FK-520

Example 5

Replacement of Methoxyl with Ethyl at C-13 and/or C-15 of FK-506 and FK-520

The present invention also provides novel FK-506 and FK-520 derivative compounds in which the methoxy groups at either or both the C-13 and C-15 positions are instead ethyl groups. These compounds are produced by novel PKS enzymes of the invention in which the AT domains of modules 8 and/or 7 are converted to ethylmalonyl specific AT domains by modification of the PKS gene that encodes the module. Ethylmalonyl specific AT domain coding sequences can be obtained from, for example, the FK-520 PKS genes, the niddamycin PKS genes, and the tylosin PKS genes. The novel PKS genes of the invention include not only those in which either or both of the AT domains of modules 7 and 8 have been converted to ethylmalonyl specific AT domain and the other is converted to a malonyl specific or a methylmalonyl specific AT domain.

5

10

15

20

25

30

35

Example 6

Neurotrophic Compounds

The compounds described in Examples 1 - 4, inclusive have immunosuppressant activity and can be employed as immunosuppressants in a manner and in formulations similar to those employed for FK-506. The compounds of the invention are generally effective for the prevention of organ rejection in patients receiving organ transplants and in particular can be used for immunosuppression following orthotopic liver transplantation. These compounds also have pharmacokinetic properties and metabolism that are more advantageous for certain applications relative to those of FK-506 or FK-520. These compounds are also neurotrophic; however, for use as neurotrophins, it is desirable to modify the compounds to diminish or abolish their immunosuppressant activity. This can be readily accomplished by hydroxylating the compounds at the C-18 position using established chemical methodology or novel FK-520 PKS genes provided by the present invention.

Thus, in one aspect, the present invention provides a method for stimulating nerve growth that comprises administering a therapeutically effective dose of 18-hydroxy-FK-520. In another embodiment, the compound administered is a C-18,20-dihydroxy-FK-520 derivative. In another embodiment, the compound administered is a C-13-desmethoxy and/or C-15-desmethoxy 18-hydroxy-FK-520 derivative. In another embodiment, the compound administered is a C-13-desmethoxy and/or C-15-desmethoxy 18,20-dihydroxy-FK-520 derivative. In other embodiments, the compounds are the corresponding analogs of FK-506. The 18-hydroxy compounds of the invention

can be prepared chemically, as described in U.S. Patent No. 5,189,042, incorporated herein by reference, or by fermentation of a recombinant host cell provided by the present invention that expresses a recombinant PKS in which the module 5 DH domain has been deleted or rendered non-functional.

5

10

15

20

25

30

The chemical methodology is as follows. A compound of the invention (~200 mg) is dissolved in 3 mL of dry methylene chloride and added to 45 μL of 2,6-lutidine, and the mixture stirred at room temperature. After 10 minutes, tert-butyldimethylsilyl trifluoromethanesulfonate (64 µL) is added by syringe. After 15 minutes, the reaction mixture is diluted with ethyl acetate, washed with saturated bicarbonate, washed with brine, and the organic phase dried over magnesium sulfate. Removal of solvent in vacuo and flash chromatography on silica gel (ethyl acetate: hexane (1:2) plus 1% methanol) gives the protected compound, which is dissolved in 95% ethanol (2.2 mL) and to which is added 53 uL of pyridine, followed by selenium dioxide (58 mg). The flask is fitted with a water condenser and heated to 70°C on a mantle. After 20 hours, the mixture is cooled to room temperature, filtered through diatomaceous earth, and the filtrate poured into a saturated sodium bicarbonate solution. This is extracted with ethyl acetate, and the organic phase is washed with brine and dried over magnesium sulfate. The solution is concentrated and purified by flash chromatography on silica gel (ethyl acetate: hexane (1:2) plus 1% methanol) to give the protected 18-hydroxy compound. This compound is dissolved in acetonitrile and treated with aqueous HF to remove the protecting groups. After dilution with ethyl acetate, the mixture is washed with saturated bicarbonate and brine, dried over magnesium sulfate, filtered, and evaporated to yield the 18-hydroxy compound. Thus, the present invention provides the C-1S-hydroxyl derivatives of the compounds described in Examples 1 - 4.

Those of skill in the art will recognize that other suitable chemical procedures can be used to prepare the novel 18-hydroxy compounds of the invention. See, e.g., Kawai *et al.*, Jan. 1993, Structure-activity profiles of macrolactam immunosuppressant FK-506 analogues, *FEBS Letters 316*(2): 107-113, incorporated herein by reference. These methods can be used to prepare both the C18-[S]-OH and C18-[R]-OH enantiomers, with the *R* enantiomer showing a somewhat lower IC₅₀, which may be preferred in some applications. See Kawai *et al.*, *supra*. Another preferred protocol is described in Umbreit and Sharpless, 1977, JACS 99(16): 1526-28, although it may be preferable to use 30 equivalents each of SeO₂ and t-BuOOH rather than the 0.02 and 3-4 equivalents, respectively, described in that reference.

All scientific and patent publications referenced herein are hereby incorporated by reference. The invention having now been described by way of written description and example, those of skill in the art will recognize that the invention can be practiced in a variety of embodiments, that the foregoing description and example is for purposes of illustration and not limitation of the following claims.

Claims

1. An isolated nucleic acid that encodes a CoA ligase, a non-ribosomal peptide synthetase, or a domain of an extender module of a polyketide synthase enzyme that synthesizes FK-520.

5

WO 00/20601

2. The isolated nucleic acid of claim 1 that encodes an extender module, said module comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.

10

3. The isolated nucleic acid of claim 1 that encodes an open reading frame said open reading frame comprising coding sequences for two or more extender modules. each extender module comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.

15

4. The isolated nucleic acid of claim 1 that encodes a gene cluster, said gene cluster comprising two or more open reading frames, each of said open reading frames comprising coding sequences for two or more extender modules, each of said extender modules comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.

20

5. The isolated nucleic acid of claim 2, wherein at least one of said domains is a domain of a module of a non-FK-520 polyketide synthase.

- 6. The isolated nucleic acid of claim 1, wherein said nucleic acid is a recombinant vector capable of replication in or integration into the chromosome of a host cell.
- 7. The isolated nucleic acid of claim 6 that is selected from the group consisting of cosmid pKOS034-120, cosmid pKOS034-124, cosmid pKOS065-M27, and cosmid 30 pKOS065-M21.
 - 8. The isolated nucleic acid of claim 5, wherein said non-FK-520 polyketide synthase is rapamycin polyketide synthase, FK-506 polyketide synthase, or erythromcyin polyketide synthase.

9. A method of preparing a polyketide, said method comprising transforming a host cell with a recombinant DNA vector of claim 6, and culturing said host cell under conditions such that said polyketide synthase is produced and catalyzes synthesis of said polyketide.

5

- 10. The method of claim 9, wherein said host cell is a Streptomyces host cell.
- 11. The method of claim 9, wherein said polyketide is selected from the group consisting of FK-520, 13-desmethoxy-FK-520, and 13-desmethoxy-FK-506.

10

15

20

- 12. A recombinant host cell that expresses a recombinant polyketide synthase selected from the group consisting of: (i) an FK-520 polyketide synthase in which at least one AT domain is replaced by an AT domain of a non-FK-520 polyketide synthase; (ii) an FK-506 polyketide synthase in which at least one AT domain is replaced by an AT domain of a non-FK-506 polyketide synthase; (iii) an FK-520 polyketide synthase in which at least one DH domain has been deleted; (iv) an FK-506 polyketide synthase in which at least one DH domain has been deleted.
- 13. The recombinant host cell of claim 12 that expresses an FK-520 polyketide synthase in which an AT domain of module 8 has been replaced by an AT domain that binds malonyl CoA, methylmalonyl CoA, or ethylmalonyl CoA.
- 14. The recombinant host cell of claim 12 that expresses an FK-506 polyketide synthase in which an AT domain of module 8 has been replaced by an AT domain that binds malonyl CoA, methylmalonyl CoA, or ethylmalonyl CoA.
- 15. The recombinant host cell of claim 13, wherein a DH domain of module 5 or module 6 has been deleted.
- 30 16. The recombinant host cell of claim 14, wherein a DH domain of module 5 or module 6 has been deleted.
 - 17. A recombinant host cell that comprises recombinant genes coding for enzymes sufficient for synthesis of ethylmalonyl CoA or 2-hydroxymalonyl CoA.

18. A polyketide having the structure

- wherein, R₁ is hydrogen, methyl, ethyl, or allyl; R₂ is hydrogen or hydroxyl, provided that when R₂ is hydrogen, there is a double bond between C-20 and C-19; R₃ is hydrogen or hydroxyl; R₄ is methoxyl, hydrogen, methyl, or ethyl; and R₅ is methoxyl, hydrogen, methyl, or ethyl; but not including FK-506, FK-520, 18-hydroxy-FK-520, and 18-hydroxy-FK-506.
- 19. The polyketide of claim 18 that is 13-desmethoxy-FK-506.
 - 20. The polyketide of claim 18 that is 13-desmethoxy-18-hydroxy-FK-520.

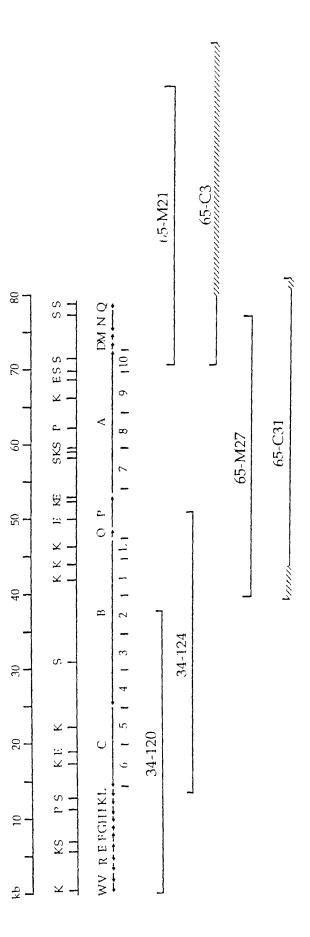


Figure 1

Figure 2

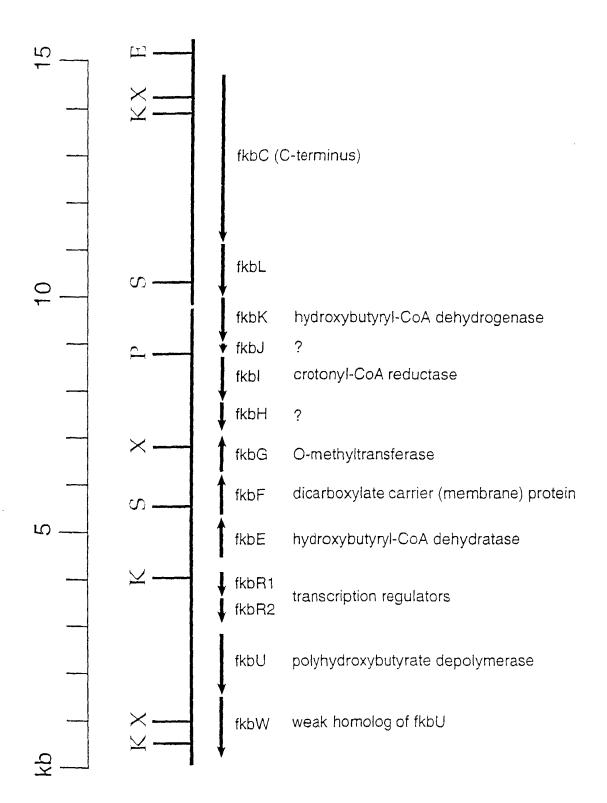


Figure 3

Figure 4

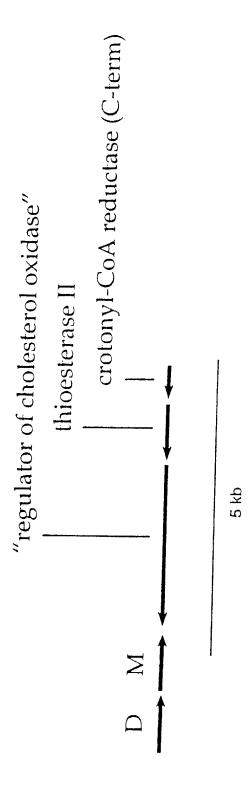


Figure 5

Figure 6

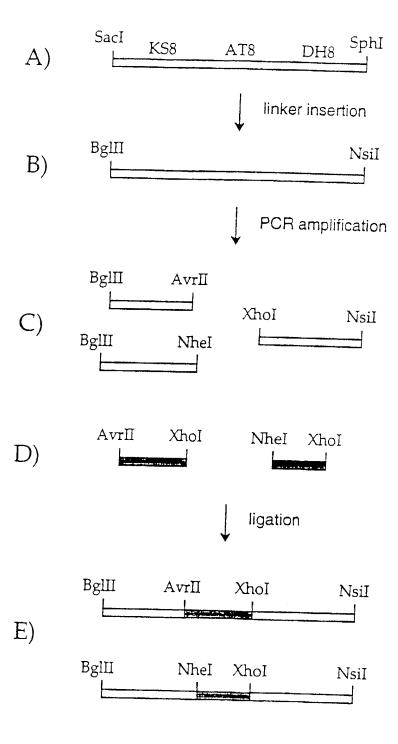


Figure 7

Figure 8 Part A



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7:

C12N 15/52, 15/54, 15/62, 9/10, C12P 17/18, 19/32, C07D 498/18 // (C07D 498/18, 311:00, 273:00, 211:00)

(11) International Publication Number:

WO 00/20601

(43) International Publication Date:

13 April 2000 (13.04.00)

(21) International Application Number:

PCT/US99/22886

A2

(22) International Filing Date:

1 October 1999 (01.10.99)

(30) Priority Data:

US 60/102,748 2 October 1998 (02.10.98) US 60/123,810 11 March 1999 (11.03.99) 60.139,650 17 June 1999 (17.06.99) US

(71) Applicant (for all designated States except US): KOSAN BIOSCIENCES, INC. [US/US]; 3832 Bay Center Drive, Hayward, CA 94545 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): REEVES, Christopher [US/US]: 4 East Altarinda Drive, Orinda, CA 94563 (US). Published CHU, Daniel [US/US]; 3767 Benton Street, Santa Clara, CA 95051 (US). KHOSLA, Chaitan [IN/US]; 740 Para Avenue, Palo Alto, CA 94306 (US), SANTI, Daniel [US/US]; 211 Belgrave Avenue, San Francisco, CA 94117 (US). WU, Kai [CN/US]: 900 Constitution Drive, Foster City, CA 94404 (US).

(74) Agents: FAVORITO, Carolyn et al.; Morrison & Foerster LLP, 2000 Pennsylvania Avenue, N.W., Washington, DC 20006-1888 (US).

(81) Designated States: AL, AM, AU, BA, BB, BG, BR, CA, CN, CR, CU, CZ, DM, EE, GD, GE, HR, HU, IL, IS, JP, KG, KP, KR, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, US, UZ, VN, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TY UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, IJ, TM). European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, R, IE, IT, LU, MC, NL, 17, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG:

Without international search report and to be republished upon receipt of that report

With an indication in relation to deposited biological material jurnished under Rule 13bis separately from the description

(54) Title: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA CONSTRUCTS THEREFOR

(57) Abstract

Host cells comprising recombinant vectors encoding the FK-520 polyketide synthase and FK-520 modification enzymes can be used to produce the FK-520 polyketide. Recombinant DNA constructs comprising one or more FK-520 polyketide synthase domains, modules, open reading frames, and variants thereof can be used to produce recombinant polyketide synthases and a variety of different polyketides with application as pharmaceutical and veterinary products.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	1.V	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	мG	Madagascar	ТJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE.	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	$\mathbf{U}\mathbf{G}$	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	us	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
cu	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	1.1	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

10

15

20

25

30

POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA CONSTRUCTS THEREFOR

Field of the Invention

The present invention relates to polyketides and the polyketide synthase (PKS) enzymes that produce them. The invention also relates generally to genes encoding PKS enzymes and to recombinant host cells containing such genes and in which expression of such genes leads to the production of polyketides. The present invention also relates to compounds useful as medicaments having immunosuppressive and/or neurotrophic activity. Thus, the invention relates to the fields of chemistry, molecular biology, and agricultural, medical, and veterinary technology.

Background of the Invention

Polyketides are a class of compounds synthesized from 2-carbon units through a series of condensations and subsequent modifications. Polyketides occur in many types of organisms, including fungi and mycelial bacteria, in particular, the actinomycetes. Polyketides are biologically active molecules with a wide variety of structures, and the class encompasses numerous compounds with diverse activities. Tetracycline, erythromycin, epothilone, FK-506, FK-520, narbomycin, picromycin, rapamycin, spinocyn, and tylosin are examples of polyketides. Given the difficulty in producing polyketide compounds by traditional chemical methodology, and the typically low production of polyketides in wild-type cells, there has been considerable interest in finding improved or alternate means to produce polyketide compounds.

This interest has resulted in the cloning, analysis, and manipulation by recombinant DNA technology of genes that encode PKS enzymes. The resulting technology allows one to manipulate a known PKS gene cluster either and roduce the polyketide synthesized by that PKS at higher levels than occur in nature or in hosts that otherwise do not produce the polyketide. The technology also allows one to produce molecules that are structurally related to, but distinct from, the polyketides produced from known PKS gene clusters. See, e.g., PCT publication Nos. WO 93/13663; 95/08548; 96/40968; 97/02358; 98/27203; and 98/49315; United States Patent Nos. 4,874,748; 5.063,155; 5.098,837; 5,149,639; 5,672,491; 5,712,146; 5,830,750; and 5.843,718; and Fu et al., 1994, Biochemistry 33:

1Ú

15

20

25

30

9321-9326; McDaniel *et al.*, 1993, *Science 262*: 1546-1550; and Rohr, 1995, *Angew. Chem. Int. Ed. Engl. 34*(8): 881-888, each of which is incorporated herein by reference.

Polyketides are synthesized in nature by PKS enzymes. These enzymes, which are complexes of multiple large proteins, are similar to the synthases that catalyze condensation of 2-carbon units in the biosynthesis of fatty acids. PKSs catalyze the biosynthesis of polyketides through repeated, decarboxylative Claisen condensations between acylthioester building blocks. The building blocks used to form complex polyketides are typically acylthioesters, such as acetyl, butyryl, propionyl, malonyl, hydroxymalonyl, methylmalonyl, and ethylmalonyl CoA. Other building blocks include amino acid like acylthioesters. PKS enzymes that incorporate such building blocks include an activity that functions as an amino acid ligase (an AMP ligase) or as a non-ribosomal peptide synthetase (NRPS). Two major types of PKS enzymes are known: these differ in their composition and mode of synthesis of the polyketide synthesized. These two major types of PKS enzymes are commonly referred to as Type I or "modular" and Type II "iterative" PKS enzymes.

In the Type I or modular PKS enzyme group, a set of separate catalytic active sites (each active site is termed a "domain", and a set thereof is termed a "module") exists for each cycle of carbon chain elongation and modification in the polyketide synthesis pathway. The typical modular PKS is composed of several large polypeptides, which can be segregated from amino to carboxy termini into a loading module, multiple extender modules, and a releasing (or thioesterase) domain. The PKS enzyme known as 6-deoxyerythronolide B synthase (DEBS) is a Type I PKS. In DEBS, there is a loading module, six extender modules, and a thioesterase (TE) domain. The loading module, six extender modules, and TE of DEBS are present on three separate proteins (designated DEBS-1, DEBS-2, and DEBS-3, with two extender modules per protein). Each of the DEBS polypeptides is encoded by a separate open reading frame (ORF) or gene; these genes are known as eryAI, eryAII, and eryAIII. See Caffrey et al., 1992, FEBS Letters 304: 205, and U.S. Patent No. 5,824,513, each of which is incorporated herein by reference.

Generally, the loading module is responsible for binding the first building block used to synthesize the polyketide and transferring it to the first extender module. The loading module of DEBS consists of an acyltransferase (AT) domain and an acyl carrier protein (ACP) domain. Another type of loading module utilizes an inactivated ketosynthase (KS) domain and AT and ACP domains. This inactivated KS is in some instances called KS^Q, where the superscript letter is the abbreviation for the amino acid, glutamine, that is

10

15

20

25

30

3

present instead of the active site cysteine required for ketosynthase activity. In other PKS enzymes, including the FK-506 PKS, the loading module incorporates an unusual starter unit and is composed of a CoA ligase like activity domain. In any event, the loading module recognizes a particular acyl-CoA (usually acetyl or propionyl but sometimes butyryl or other acyl-CoA) and transfers it as a thiol ester to the ACP of the loading module.

The AT on each of the extender modules recognizes a particular extender-CoA (malonyl or alpha-substituted malonyl, i.e., methylmalonyl, ethylmalonyl, and 2-hydroxymalonyl) and transfers it to the ACP of that extender module to form a thioester. Each extender module is responsible for accepting a compound from a prior module, binding a building block, attaching the building block to the compound from the prior module, optionally performing one or more additional functions, and transferring the resulting compound to the next module.

Each extender module of a modular PKS contains a KS, AT, ACP, and zero, one, two, or three domains that modify the beta-carbon of the growing polyketide chain. A typical (non-loading) minimal Type I PKS extender module is exemplified by extender module three of DEBS, which contains a KS domain, an AT domain, and an ACP domain. These three domains are sufficient to activate a 2-carbon extender unit and attach it to the growing polyketide molecule. The next extender module, in turn, is responsible for attaching the next building block and transferring the growing compound to the next extender module until synthesis is complete.

Once the PKS is primed with acyl- and malonyl-ACPs, the acyl group of the loading module is transferred to form a thiol ester (trans-esterification) at the KS of the first extender module; at this stage, extender module one possesses an acyl-KS and a malonyl (or substituted malonyl) ACP. The acyl group derived from the loading module is then covalently attached to the alpha-carbon of the malonyl group to form a carbon-carbon bond, driven by concomitant decarboxylation, and generating a new acyl-ACP that has a backbone two carbons longer than the loading building block (elongation or extension).

The polyketide chain, growing by two carbons each extender module, is sequentially passed as covalently bound thiol esters from extender module to extender module, in an assembly line-like process. The carbon chain produced by this process alone would possess a ketone at every other carbon atom, producing a polyketone, from which the name polyketide arises. Most commonly, however, additional enzymatic activities modify the beta

10

15

20

25

30

keto group of each two carbon unit just after it has been added to the growing polyketide chain but before it is transferred to the next module.

Thus, in addition to the minimal module containing KS. AT, and ACP domains necessary to form the carbon-carbon bond, and as noted above, other domains that modify the beta-carbonyl moiety can be present. Thus, modules may contain a ketoreductase (KR) domain that reduces the keto group to an alcohol. Modules may also contain a KR domain plus a dehydratase (DH) domain that dehydrates the alcohol to a double bond. Modules may also contain a KR domain, and an enoylreductase (ER) domain that converts the double bond product to a saturated single bond using the beta carbon as a methylene function. An extender module can also contain other enzymatic activities, such as, for example, a methylase or dimethylase activity.

After traversing the final extender module, the polyketide encounters a releasing domain that cleaves the polyketide from the PKS and typically cyclizes the polyketide. For example, final synthesis of 6-dEB is regulated by a TE domain located at the end of extender module six. In the synthesis of 6-dEB, the TE domain catalyzes cyclization of the macrolide ring by formation of an ester linkage. In FK-506, FK-520, rapamycin, and similar polyketides, the TE activity is replaced by a RapP (for rapamycin) or RapP like activity that makes a linkage incorporating a pipecolate acid residue. The enzymatic activity that catalyzes this incorporation for the rapamycin enzyme is known as RapP, encoded by the *rapP* gene. The polyketide can be modified further by tailoring enzymes: these enzymes add carbohydrate groups or methyl groups, or make other modifications, i.e., oxidation or reduction, on the polyketide core molecule. For example, 6-dEB is hydroxylated at C-6 and C-12 and glycosylated at C-3 and C-5 in the synthesis of erythromycin A.

In Type I PKS polypeptides, the order of catalytic domains is conserved. When all beta-keto processing domains are present in a module, the order of domains in that module from N-to-C-terminus is always KS, AT, DH, ER, KR, and ACP. Some or all of the beta-keto processing domains may be missing in particular modules, but the order of the domains present in a module remains the same. The order of domains within modules is believed to be important for proper folding of the PKS polypetides into an active complex. Importantly, there is considerable flexibility in PKS enzymes, which allows for the genetic engineering of novel catalytic complexes. The engineering of these enzymes is achieved by modifying, adding, or deleting domains, or replacing them with those taken from other Type I PKS enzymes. It is also achieved by deleting, replacing, or adding entire modules with those

10

15

20

25

30

taken from other sources. A genetically engineered PKS complex should of course have the ability to catalyze the synthesis of the product predicted from the genetic alterations made.

Alignments of the many available amino acid sequences for Type I PKS enzymes has approximately defined the boundaries of the various catalytic domains. Sequence alignments also have revealed linker regions between the catalytic domains and at the Nand C-termini of individual polypeptides. The sequences of these linker regions are less well conserved than are those for the catalytic domains, which is in part how linker regions are identified. Linker regions can be important for proper association between domains and agual polypeptides that comprise the PWS complex. One can thus view the between the linkers and domains together as creating a scaffold on which the domains and modules are positioned in the correct orientation to be active. This organization and positioning, if retained, permits PKS domains of different or identical substrate specificities to be substituted (usually at the DNA level) between PKS enzymes by various available methodologies. In selecting the boundaries of, for example, an AT replacement, one can thus make the replacement so as to retain the linkers of the recipient PKS or to replace them with the linkers of the donor PKS AT domain, or, preferably, make both constructs to ensure that the correct linker regions between the KS and AT domains have been included in at least one of the engineered enzymes. Thus, there is considerable flexibility in the design of new PKS enzymes with the result that known polyketides can be produced more effectively, and novel polyketides useful as pharmaceuticals or for other purposes can be made.

By appropriate application of recombinant DNA technology, a wide variety of polyketides can be prepared in a variety of different host cells provided one has access to nucleic acid compounds that encode PKS proteins and polyketide modification enzymes. The present invention helps meet the need for such nucleic acid compounds by providing recombinant vectors that encode the FK-520 PKS enzyme and various FK-520 modification enzymes. Moreover, while the FK-506 and FK-520 polyketides have many useful activities, there remains a need for compounds with similar useful activities but with better pharmacokinetic profile and metabolism and fewer side-effects. The present invention helps meet the need for such compounds as well.

Summary of the Invention

10

15

20

25

30

In one embodiment, the present invention provides recombinant DNA vectors that encode all or part of the FK-520 PKS enzyme. Illustrative vectors of the invention include cosmid pKOS034-120, pKOS034-124, pKOS065-C31, pKOS065-C3, pKOS065-M27, and pKOS065-M21. The invention also provides nucleic acid compounds that encode the various domains of the FK-520 PKS, i.e., the KS, AT, ACP, KR, DH, and ER domains. These compounds can be readily used, alone or in combination with nucleic acids encoding other FK-520 or non-FK-520 PKS domains, as intermediates in the construction of recombinant vectors that encode all or part of PKS enzymes that make novel polyketides.

The invention also provides isolated nucleic acids that encode all or part of one or more modules of the FK-520 PKS, each module comprising a ketosynthase activity, an acyl transferase activity, and an acyl carrier protein activity. The invention provides an isolated nucleic acid that encodes one or more open reading frames of FK-520 PKS genes, said open reading frames comprising coding sequences for a CoA ligase activity, an NRPS activity, or two or more extender modules. The invention also provides recombinant expression vectors containing these nucleic acids.

In another embodiment, the invention provides isolated nucleic acids that encode all or a part of a PKS that contains at least one module in which at least one of the domains in the module is a domain from a non-FK-520 PKS and at least one domain is from the FK-520 PKS. The non-FK-520 PKS domain or module originates from the rapamycin PKS, the FK-506 PKS, DEBS, or another PKS. The invention also provides recombinant expression vectors containing these nucleic acids.

In another embodiment, the invention provides a method of preparing a polyketide, said method comprising transforming a host cell with a recombinant DNA vector that encodes at least one module of a PKS, said module comprising at least one FK-520 PKS domain, and culturing said host cell under conditions such that said PKS is produced and catalyzes synthesis of said polyketide. In one aspect, the method is practiced with a *Streptomyces* host cell. In another aspect, the polyketide produced is FK-520. In another aspect, the polyketide produced is a polyketide related in structure to FK-520 or rapamycin.

In another embodiment, the invention provides a set of genes in recombinant form sufficient for the synthesis of ethylmalonyl CoA in a heterologous host cell. These genes and the methods of the invention enable one to create recombinant host cells with the ability to produce polyketides or other compounds that require ethylmalonyl CoA for biosynthesis.

The invention also provides recombinant nucleic acids that encode AT domains specific for ethylmalonyl CoA. Thus, the compounds of the invention can be used to produce polyketides requiring ethylmalonyl CoA in host cells that otherwise are unable to produce such polyketides.

In another embodiment, the invention provides a set of genes in recombinant form sufficient for the synthesis of 2-hydroxymalonyl CoA and 2-methoxymalonyl CoA in a heterologous host cell. These genes and the methods of the invention enable one to create recombinant host cells with the ability to produce polyketides or other compounds that require 2-hydroxymalonyl CoA for biosynth siz. The invention also provides recombinant nucleic acids that encode AT domains specific for 2-hydroxymalonyl CoA and 2-methoxymalonyl CoA. Thus, the compounds of the invention can be used to produce polyketides requiring 2-hydroxymalonyl CoA or 2-methoxymalonyl CoA in host cells that are otherwise unable to produce such polyketides.

In another embodiment, the invention provides a compound related in structure to FK-520 or FK-506 that is useful in the treatment of a medical condition. These compounds include compounds in which the C-13 methoxy group is replaced by a moiety selected from the group consisting of hydrogen, methyl, and ethyl moieties. Such compounds are less susceptible to the main *in vivo* pathway of degradation for FK-520 and FK-506 and related compounds and thus exhibit an improved pharmacokinetic profile. The compounds of the invention also include compounds in which the C-15 methoxy group is replaced by a moiety selected from the group consisting of hydrogen, methyl, and ethyl moieties. The compounds of the invention also include the above compounds further modified by chemical methodology to produce derivatives such as, but not limited to, the C-18 hydroxyl derivatives, which have potent neurotrophin but not immunosuppresion activities.

Thus, the invention provides polyketides having the structure:

5

10

15

20

wherein, R₁ is hydrogen, methyl, or allyl; R₂ is hydrogen or hydroxyl, provided that when R₂ is hydrogen, there is a double bond between C-20 and C-19; R₃ is hydrogen or hydroxyl; R₄ is methoxyl, hydrogen, methyl, or ethyl; and R₅ is methoxyl, hydrogen, methyl, or ethyl; but not including FK-506, FK-520, 18-hydroxy-FK-520, and 18-hydroxy-FK-506. The invention provides these compounds in purified form and in pharmaceutical compositions.

In another embodiment, the invention provides a method for treating a medical condition by administering a pharmaceutically efficacious dose of a compound of the invention. The compounds of the invention may be administered to achieve immunosuppression or to stimulate nerve growth and regeneration.

These and other embodiments and aspects of the invention will be more fully understood after consideration of the attached Drawings and their brief description below, together with the detailed description, examples, and claims that follow.

15

20

25

30

Brief Description of the Drawings

Figure 1 shows a diagram of the FK-520 biosynthetic gene cluster. The top line provides a scale in kilobase pairs (kb). The second line shows a restriction map with selected restriction enzyme recognition sequences indicated. K is *KpnI*; X is *XhoI*, S is *SacI*; P is *PstI*; and E is *EcoRI*. The third line indicates the position of FK-520 PKS and related genes. Genes are abbreviated with a one letter designation, i.e., C is *fkbC*. Immediately under the third line are numbered segments showing where the loading module (L) and ten different extender modules (numbered 1 - 10) are encoded on the various genes shown. At the bottom of the Figure, the DNA inserts of various clamids of the invention (i.e., 34-124 is cosmid pKOS034-124) are shown in alignment with the FK-520 biosynthetic gene cluster.

Figure 2 shows the loading module (load), the ten extender modules, and the peptide synthetase domain of the FK-520 PKS, together with, on the top line, the genes that encode the various domains and modules. Also shown are the various intermediates in FK-520 biosynthesis, as well as the structure of FK-520, with carbons 13, 15, 21, and 31 numbered. The various domains of each module and subdomains of the loading module are also shown. The darkened circles showing the DH domains in modules 2, 3, and 4 indicate that the dehydratase domain is not functional as a dehydratase; this domain may affect the

10

15

20

25

30

stereochemistry at the corresponding position in the polyketide. The substituents on the FK-520 structure that result from the action of non-PKS enzymes are also indicated by arrows, together with the types of enzymes or the genes that code for the enzymes that mediate the action. Although the methyltransferase is shown acting at the C-13 and C-15 hydroxyl groups after release of the polyketide from the PKS, the methyltransferase may act on the 2-hydroxymalonyl substrate prior to or contemporaneously with its incorporation during polyketide synthesis.

Figure 3 shows a close-up view of the left end of the FK-520 gene cluster, which contains at least ten additional genes. The ethyl side chair, on carbon 21 of FK-520 (Figure 2) is derived from an ethylmalonyl CoA extender unit that is incorporated by an ethylmalonyl specific AT domain in extender module 4 of the PKS. At least four of the genes in this region code for enzymes involved in ethylmalonyl biosynthesis. The polyhydroxybutyrate depolymerase is involved in maintaining hydroxybutyryl-CoA pools during FK-520 production. Polyhydroxybutyrate accumulates during vegetative growth and disappears during stationary phase in other *Streptomyces* (Ranade and Vining, 1993, *Can. J. Microbiol.* 39:377). Open reading frames with unknown function are indicated with a question mark.

Figure 4 shows a biosynthetic pathway for the biosynthesis of ethylmalonyl CoA from acetoacetyl CoA consistent with the function assigned to four of the genes in the FK-520 gene cluster shown in Figure 3.

Figure 5 shows a close-up view of the right-end of the FK-520 PKS gene cluster (and of the sequences on cosmid pKOS065-C31). The genes shown include *fkbD*. *fkbM* (a methyl transferase that methylates the hydroxyl group on C-31 of FK-520), *fkbN* (a homolog of a gene described as a regulator of cholesterol oxidase and that is believed to be a transcriptional activator), *fkbQ* (a type II thioesterase, which can increase polyketide production levels), and *fkbS* (a crotonyl-CoA reductase involved in the biosynthesis of ethylmalonyl CoA).

Figure 6 shows the proposed degradative pathway for tacrolimus (FK-506) metabolism.

Figure 7 shows a schematic process for the construction of recombinant PKS genes of the invention that encode PKS enzymes that produce 13-desmethoxy FK-506 and FK-520 polyketides of the invention, as described in Example 4, below.

10

15

20

Figure 8, in Parts A and B, shows certain compounds of the invention preferred for dermal application in Part A and a synthetic route for making those compounds in Part B.

Detailed Description of the Invention

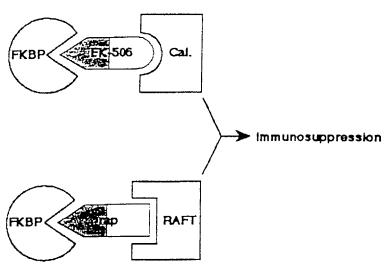
Given the valuable pharmaceutical properties of polyketides, there is a need for methods and reagents for producing large quantities of polyketides, as well as for producing related compounds not found in nature. The present invention provides such methods and reagents, with particular application to methods and reagents for producing the polyketides known as FK-520, also known as ascomycin or L-683,590 (see ! lit et al., 1993, JACS 115:9925), and FK-506, also known as tacrolimus. Tacrolimus is a macrolide immunosuppressant used to prevent or treat rejection of transplanted heart, kidney, liver, lung, pancreas, and small bowel allografts. The drug is also useful for the prevention and treatment of graft-versus-host disease in patients receiving bone marrow transplants, and for the treatment of severe, refractory uveitis. There have been additional reports of the unapproved use of tacrolimus for other conditions, including alopecia universalis, autoimmune chronic active hepatitis, inflammatory bowel disease, multiple sclerosis, primary biliary cirrhosis, and scleroderma. The invention provides methods and reagents for making novel polyketides related in structure to FK-520 and FK-506, and structurally related polyketides such as rapamycin.

The FK-506 and rapamycin polyketides are potent immunosuppressants, with chemical structures shown below.

10

FK-520 differs from FK-506 in that it lacks the allyl group at C-21 of FK-506, having instead an ethyl group at that position, and has similar activity to FK-506, albeit reduced immunosuppressive activity.

These compounds act through initial formation of an intermediate complex with protein "immunophilins" known as FKBPs (FK-506 binding proteins), including FKBP-12. Immunophilins are a class of cytosolic proteins that form complexes with molecules such as FK-506, FK-520, and rapamycin that in turn serve as ligands for other cellular targets involved in signal transduction. Binding of FK-506, FK-520, and rapamycin to FKBP occurs through the structurally similar segments of the polyketide molecules, known as the "FKBP-binding domain" (as generally but not precisely indicated by the stippled regions in the structures above). The FK-506-FKBP complex then binds calcincurin, while the rapamycin-FKBP complex binds to a protein known as RAFT-1. Binding of the FKBP-polyketide complex to these second proteins occurs through the dissimilar regions of the drugs known as the "effector" domains.



15

20

The three component FKBP-polyketide-effector complex is required for signal transduction and subsequent immunosuppressive activity of FK-506. FK-520, and rapamycin. Modifications in the effector domains of FK-506. FK-520, and rapamycin that destroy binding to the effector proteins (calcineurin or RAFT) lead to loss of immunosuppressive activity, even though FKBP binding is unaffected. Further, such analogs antagonize the immunosuppressive effects of the parent polyketides, because they compete for FKBP. Such non-immunosuppressive analogs also show reduced toxicity (see Dumont et al., 1992, Journal of Experimental Medicine 176, 751-760), indicating that much of the toxicity of these drugs is not linked to FKBP binding.

}()

15

20

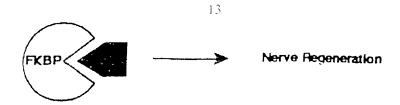
25

30

In addition to immunosuppressive activity. FK-520. FK-506, and rapamycin have neurotrophic activity. In the central nervous system and in peripheral nerves, immunophilins are referred to as "neuroimmunophilins". The neuroimmunophilin FKBP is markedly enriched in the central nervous system and in peripheral nerves. Molecules that bind to the neuroimmunophilin FKBP, such as FK-506 and FK-520, have the remarkable effect of stimulating nerve growth. *In vitro*, they act as neurotrophins, i.e., they promote neurite outgrowth in NGF-treated PC12 cells and in sensory neuronal cultures, and in intact animals, they promote regrowth of damaged facial and sciatic nerves, and repair lesioned serotonin and dopamine neurons in the brain. See Gold *et al.*, Jun. 1999, *J. Pharm. Exp. Ther.* 289(3): 1202-1210; Lyons *et al.*, 1994, *Proc. National Academy of Science* 91: 3191-3195; Gold *et al.*, 1995, *Journal of Neuroscience* 15: 7509-7516; and Steiner *et al.*, 1997, *Proc. National Academy of Science* 94: 2019-2024. Further, the restored central and peripheral neurons appear to be functional.

Compared to protein neurotrophic molecules (BNDF, NGF, etc.), the small-molecule neurotrophins such as FK-506, FK-520, and rapamycin have different, and often advantageous, properties. First, whereas protein neurotrophins are difficult to deliver to their intended site of action and may require intra-cranial injection, the small-molecule neurotrophins display excellent bioavailability; they are active when administered subcutaneously and orally. Second, whereas protein neurotrophins show quite specific effects, the small-molecule neurotrophins show rather broad effects. Finally, whereas protein neurotrophins often show effects on normal sensory nerves, the small-molecule neurotrophins do not induce aberrant sprouting of normal neuronal processes and seem to affect damaged nerves specifically. Neuroimmunophilin ligands have potential therapeutic utility in a variety of disorders involving nerve degeneration (e.g. multiple sclerosis, Parkinson's disease, Alzheimer's disease, stroke, traumatic spinal cord and brain injury, peripheral neuropathies).

Recent studies have shown that the immunosuppressive and neurite outgrowth activity of FK-506, FK-520, and rapamycin can be separated; the neuroregenerative activity in the absence of immunosuppressive activity is retained by agents which bind to FKBP but not to the effector proteins calcineurin or RAFT. See Steiner *et al.*, 1997, *Nature Medicine* 3: 421-428.



Available structure-activity data show that the important features for neurotrophic activity of rapamycin, FK-520, and FK-506 lie within the common, contiguous segments of the macrolide ring that bind to FKBP. This portion of the molecule is termed the "FKBP binding domain" (see VanDuyne et al., 1993, Journal of Molecular Biology 229: 105-124.). Nevertheless, the effector domains of the parent macrolides contribute to conformational rigidity of the binding domain and thus indirectly contribute to FKBP binding.

5

10

15

"FKBP binding domain"

There are a number of other reported analogs of FK-506, FK-520, and rapamycin that bind to FKBP but not the effector protein calcineurin or RAFT. These analogs show effects on nerve regeneration without immunosuppressive effects.

Naturally occurring FK-520 and FK-506 analogs include the antascomycins, which are FK-506-like macrolides that lack the functional groups of FK-506 that bind to calcineurin (see Fehr et al., 1996, *The Journal of Antibiotics 49*: 230-233). These molecules bind FKBP as effectively as does FK-506; they antagonize the effects of both FK-506 and rapamycin, yet lack immunosuppressive activity.

10

15

Antascomycin A

Other analogs can be produced by chemically modifying FK-506, FK-520, or rapamycin. One approach to obtaining neuroimmunophilin ligands is to destroy the effector binding region of FK-506, FK-520, or rapamycin by chemical modification. While the chemical modifications permitted on the parent compounds are quite limited, some useful chemically modified analogs exist. The FK-520 analog L-685,818 (ED₅₀ = 0.7 nM for FKBP binding; see Dumont *et al.*, 1992), and the rapamycin analog WAY-124,466 (IC₅₀ = 12.5 nM; see Ocain *et al.*, 1993, *Biochemistry Biophysical Research Communications 192*: 1340-134693) are about as effective as FK-506, FK-520, and rapamycin at promoting neurite outgrowth in sensory neurons (see Steiner *et al.*, 1997).

One of the few positions of rapamycin that is readily amenable to chemical modification is the allylic 16-methoxy group; this reactive group is readily exchanged by acid-catalyzed nucleophilic substitution. Replacement of the 16-methoxy group of rapamycin with a variety of bulky groups has produced analogs showing selective loss of immunosuppressive activity while retaining FKBP-binding (see Luengo *et al.*, 1995, *Chemistry & Biology 2*: 471-481). One of the best compounds, 1, below, shows complete

loss of activity in the splenocyte proliferation assay with only a 10-fold reduction in binding to FKBP.

There are also synthetic analogs of FKBP binding domains. These compounds reflect an approach to obtaining neuroimmunophilin ligands based on "rationally designed" molecules that retain the FKBP-binding region in an appropriate conformation for binding to FKBP, but do not possess the effector binding regions. In one example, the ends of the FKBP binding domain were tethered by hydrocarbon chains (see Holt *et al.*, 1993, *Journal of the American Chemical Society 115*: 9925-9938); the best analog, **2**, below, binds to FKBP about as well as FK-506. In a similar approach, the ends of the FKBP binding domain were tethered by a tripeptide to give analog **3**, below, which binds to FKBP about 20-fold poorer than FK-506. These compounds are anticipated to have neuroimmunophilin binding activity.

15

20

5

10

In a primate MPTP model of Parkinson's disease, administration of FKBP ligand GPI-1046 caused brain cells to regenerate and behavioral measures to improve. MPTP is a neurotoxin, which, when administered to animals, selectively damages nigral-striatal dopamine neurons in the brain, mimicking the damage caused by Parkinson's disease. Whereas, before treatment, animals were unable to use affected limbs, the FKBP ligand

10

15

20

25

30

restored the ability of animals to feed themselves and gave improvements in measures of locomotor activity, neurological outcome, and fine motor control. There were also corresponding increases in regrowth of damaged nerve terminals. These results demonstrate the utility of FKBP ligands for treatment of diseases of the CNS.

From the above description, two general approaches towards the design of non-immunosuppressant, neuroimmunophilin ligands can be seen. The first involves the construction of constrained cyclic analogs of FK-506 in which the FKBP binding domain is fixed in a conformation optimal for binding to FKBP. The advantages of this approach are that the conformation of the analogs can be accurately modeled and predicted by computational methods, and the analogs closely resemble parent molecules that have proven pharmacological properties. A disadvantage is that the difficult chemistry limits the numbers and types of compounds that can be prepared. The second approach involves the trial and error construction of acyclic analogs of the FKBP binding domain by conventional medicinal chemistry. The advantages to this approach are that the chemistry is suitable for production of the numerous compounds needed for such interactive chemistry-bioassay approaches. The disadvantages are that the molecular types of compounds that have emerged have no known history of appropriate pharmacological properties, have rather labile ester functional groups, and are too conformationally mobile to allow accurate prediction of conformational properties.

The present invention provides useful methods and reagents related to the first approach, but with significant advantages. The invention provides recombinant PKS genes that produce a wide variety of polyketides that cannot otherwise be readily synthesized by chemical methodology alone. Moreover, the present invention provides polyketides that have either or both of the desired immunosuppressive and neurotrophic activities, some of which are produced only by fermentation and others of which are produced by fermentation and chemical modification. Thus, in one aspect, the invention provides compounds that optimally bind to FKBP but do not bind to the effector proteins. The methods and reagents of the invention can be used to prepare numerous constrained cyclic analogs of FK-520 in which the FKBP binding domain is fixed in a conformation optimal for binding to FKBP. Such compounds will show neuroimmunophilin binding (neurotrophic) but not immunosuppressive effects. The invention also allows direct manipulation of FK-520 and related chemical structures via genetic engineering of the enzymes involved in the biosynthesis of FK-520 (as well as related compounds, such as FK-506 and rapamycin);

15

20

25

30

similar chemical modifications are simply not possible because of the complexity of the structures. The invention can also be used to introduce "chemical handles" into normally inert positions that permit subsequent chemical modifications.

Several general approaches to achieve the development of novel neuroimmunophilin ligands are facilitated by the methods and reagents of the present invention. One approach is to make "point mutations" of the functional groups of the parent FK-520 structure that bind to the effector molecules to eliminate their binding potential. These types of structural modifications are difficult to perform by chemical modification, but can be readily accomplished with the methods and reagents of the invention.

A second, more extensive approach facilitated by the present invention is to utilize molecular modeling to predict optimal structures *ab muto* that bind to FKBP but not effector molecules. Using the available X-ray crystal structure of FK-520 (or FK-506) bound to FKBP, molecular modeling can be used to predict polyketides that should optimally bind to FKBP but not calcineurin. Various macrolide structures can be generated by linking the ends of the FKBP-binding domain with "all possible" polyketide chains of variable length and substitution patterns that can be prepared by genetic manipulation of the FK-520 or FK-506 PKS gene cluster in accordance with the methods of the invention. The ground state conformations of the virtual library can be determined, and compounds that possess binding domains most likely to bind well to FKBP can be prepared and tested.

Once a compound is identified in accordance with the above approaches, the invention can be used to generate a focused library of analogs around the lead candidate, to "fine tune" the compound for optimal properties. Finally, the genetic engineering methods of the invention can be directed towards producing "chemical handles" that enable medicinal chemists to modify positions of the molecule previously inert to chemical modification. This opens the path to previously prohibited chemical optimization of lead compounds by time-proven approaches.

Moreover, the present invention provides polyketide compounds and the recombinant genes for the PKS enzymes that produce the compounds that have significant advantages over FK-506 and FK-520 and their analogs. The metabolism and pharmacokinetics of tacrolimus has been exstensively studied, and FK-520 is believed to be similar in these respects. Absorption of tacrolimus is rapid, variable, and incomplete from the gastrointestinal tract (Harrison's Principles of Internal Medicine, 14th edition, 1998, McGraw Hill, 14, 20, 21, 64-67). The mean bioavailability of the oral dosage form is 27%,

20

25

30

trange 5 to 65%. The volume of distribution (VoID) based on plasma is 5 to 65 L per kg of body weight (L kg), and is much higher than the VoID based on whole blood concentrations, the difference reflecting the binding of tacrolimus to red blood cells. Whole blood concentrations may be 12 to 67 times the plasma concentrations. Protein binding is high (75 to 99%), primarily to albumin and alpha1-acid glycoprotein. The half-life for distribution is 0.9 hour; elimination is biphasic and variable: terminal-11.3 hr (range, 3.5 to 40.5 hours). The time to peak concentration is 0.5 to 4 hours after oral administration.

Tacrolimus is metabolized primarily by cytochrome P450 3A enzymes in the liver and small intestine. The drug is extensively metabolized with less than 1% excreted unchanged in urine. Because hepatic dysfunction decreases clearance of tacrolimus, doses have to be reduced substantially in primary graft non-function, especially in children. In addition, drugs that induce the cytochrome P450 3A enzymes reduce tacrolimus levels, while drugs that inhibit these P450s increase tacrolimus levels. Tacrolimus bioavailability doubles with co-administration of ketoconazole, a drug that inhibits P450 3A. See, Vincent et al., 1992, In vitro metabolism of FK-506 in rat, rabbit, and human liver microsomes: Identification of a major metabolite and of cytochrome P450 3A as the major enzymes responsible for its metabolism, Arch. Biochem. Biophys. 294: 454-460; Iwasaki et al., 1993. Isolation, identification, and biological activities of oxidative metabolites of FK-506, a potent immunosuppressive macrolide lactone, Drug Metabolism & Disposition 21: 971-977; Shiraga et al., 1994, Metabolism of FK-506, a potent immunosuppressive agent, by cytochrome P450 3A enzymes in rat, dog, and human liver microsomes, Biochem. Pharmacol. 47: 727-735; and Iwasaki et al., 1995. Further metabolism of FK-506 (Tacrolimus); Identification and biological activities of the metabolites oxidized at multiple sites of FK-506, Drug Metabolism & Disposition 23: 28-34. The cytochrome P450 3A subfamily of isozymes has been implicated as important in this degradative process.

Structures of the eight isolated metabolites formed by liver microsomes are shown in Figure 6. Four metabolites of FK-506 involve demethylation of the oxygens on carbons 13, 15, and 31, and hydroxylation of carbon 12. The 13-demethylated (hydroxy) compounds undergo cyclizations of the 13-hydroxy at C-10 to give MI, MVI and MVII, and the 12-hydroxy metabolite at C-10 to give I. Another four metabolites formed by oxidation of the four metabolites mentioned above were isolated by liver microsomes from dexamethasone treated rats. Three of these are metabolites doubly demethylated at the methoxy groups on carbons 15 and 31 (M-V), 13 and 31 (M-VI), and 13 and 15 (M-VII). The fourth, M-VIII,

10

15

20

25

30

was the metabolite produced after demethylation of the 31-methoxy group, followed by formation of a fused ring system by further oxidation. Among the eight metabolites, M-II has immunosuppressive activity comparable to that of FK-500, whereas the other metabolites exhibit weak or negligible activities. Importantly, the major metabolite of human, dog, and rat liver microsomes is the 13-demethylated and cyclized FK-506 (M-I).

Thus, the major metabolism of FK-506 proceeds via 13-demethylation followed by cyclization to the inactive M-I, this representing about 90% of the metabolic products after a 10 minute incubation with liver microsomes. Analogs of tacrolimus that do not possess a C-13 methoxy group would not be susceptible to the first and most important biotransformation in the destructive metabolism of tacrolimus (i.e. cyclization of 13-hydroxy to C-10). Thus, a 13-desmethoxy analog of FK-500 should have a longer half-life in the body than does FK-506. The C-13 methoxy group is believed not to be required for binding to FKBP or calcineurin. The C-13 methoxy is not present on the identical position of rapamycin, which binds to FKBP with equipotent affinity as tacrolimus. Also, analysis of the 3-dimensional structure of the FKBP-tacrolimus-calcineurin complex shows that the C-13 methoxy has no interaction with FKBP and only a minor interaction with calcineurin. The present invention provides C-13-desmethoxy analogs of FK-506 and FK-520, as well as the recombinant genes that encode the PKS enzymes that catalyze their synthesis and host cells that produce the compounds.

These compounds exhibit, relative to their naturally occurring counterparts, prolonged immunosuppressive action *in vivo*, thereby allowing a lower dosage and or reduced frequency of administration. Dosing is more predictable, because the variability in FK-506 dosage is largely due to variation of metabolism rate. FK-506 levels in blood can vary widely depending on interactions with drugs that induce or inhibit cytochrome P450 3A (summarized in USP Drug Information for the Health Care Professional). Of particular importance are the numerous drugs that inhibit or compete for CYP 3A, because they increase FK-506 blood levels and lead to toxicity (Prograf package insert, Fujisawa DUS, Rev 4'97, Rec 6'97). Also important are the drugs that induce P450 3A (e.g. Dexamethasone), because they decrease FK-506 blood levels and reduce efficacy. Because the major site of CYP 3A action on FK-506 is removed in the analogs provided by the present invention, those analogs are not as susceptible to drug interactions as the naturally occurring compounds.

15

20

25

30

Hyperglycemia, nephrotoxicity, and neurotoxicity are the most significant adverse effects resulting from the use of FK-506 and are beheved to be similar for FK-520. Because these effects appear to occur primarily by the same mechanism as the immunosuppressive action (i.e. FKBP-calcineurin interaction), the intrinsic toxicity of the desmethoxy analogs may be similar to FK-506. However, toxicity of FK-506 is dose related and correlates with high blood levels of the drug (Prograf package insert, Fujisawa EUS, Rev 4/97, Rec 6/97). Because the levels of the compounds provided by the present invention should be more controllable, the incidence of toxicity should be significantly decreased with the 13-desmethox—nalogs. Some reports show that certain FK-506 metabolites are more toxic than FK-506 itself, and this provides an additional reason to expect that a CYP 3A resistant analog can have lower toxicity and a higher therapeutic index.

Thus, the present invention provides novel compounds related in structure to FK-506 and FK-520 but with improved properties. The invention also provides methods for making these compounds by fermentation of recombinant host cells, as well as the recombinant host cells, the recombinant vectors in those host cells, and the recombinant proteins encoded by those vectors. The present invention also provides other valuable materials useful in the construction of these recombinant vectors that have many other important applications as well. In particular, the present invention provides the FK-520 PKS genes, as well as certain genes involved in the biosynthesis of FK-520 in recombinant form.

FK-520 is produced at relatively low levels in the naturally occurring cells, *Streptomyces hygroscopicus* var. *ascomyceticus*, in which it was first identified. Thus, another benefit provided by the recombinant FK-520 PKS and related genes of the present invention is the ability to produce FK-520 in greater quantities in the recombinant host cells provided by the invention. The invention also provides methods for making novel FK-520 analogs, in addition to the desmethoxy analogs described above, and derivatives in recombinant host cells of any origin.

The biosynthesis of FK-520 involves the action of several enzymes. The FK-520 PKS enzyme, which is composed of the fkbA, fkbB, fkbC, and fkbP gene products, synthesizes the core structure of the molecule. There is also a hydroxylation at C-9 mediated by the P450 hydroxylase that is the fkbD gene product and that is oxidized by the fkbO gene product to result in the formation of a keto group at C-9. There is also a methylation at C-31 that is mediated by an O-methyltransferase that is the fkbM gene product. There are also methylations at the C-13 and C-15 positions by a methyltransferase believed to be encoded

10

15

20

25

30

by the fkbG gene; this methyltransferase may act on the hydroxymalonyl CoA substrates prior to binding of the substrate to the AT domains of the PKS during polyketide synthesis. The present invention provides the genes encoding these enzymes in recombinant form. The invention also provides the genes encoding the enzymes involved in ethylmalonyl CoA and 2-hydroxymalonyl CoA biosynthesis in recombinant form. Moreover, the invention provides *Streptomyces hygroscopicus* var. ascomyceticus recombinant host cells lacking one or more of these genes that are useful in the production of useful compounds.

The cells are useful in production in a variety of ways. First, certain cells make a useful i K-1 — ciated compound merely as a result of — creation of one or more of the FK-520 biosynthesis genes. Thus, by inactivating the C-31 O-methyltransferase gene in *Streptomyces hygroscopicus* var. *ascomyceticus*, one creates a host cell that makes a desmethyl (at C-31) derivative of FK-520. Second, other cells of the invention are unable to make FK-520 or FK-520 related compounds due to an inactivation of one or more of the PKS genes. These cells are useful in the production of other polyketides produced by PKS enzymes that are encoded on recombinant expression vectors and introduced into the host cell.

Moreover, if only one PKS gene is inactivated, the ability to produce FK-520 or an FK-520 derivative compound is restored by introduction of a recombinant expression vector that contains the functional gene in a modified or unmodified form. The introduced gene produces a gene product that, together with the other endogenous and functional gene products, produces the desired compound. This methodology enables one to produce FK-520 derivative compounds without requiring that all of the genes for the PKS enzyme be present on one or more expression vectors. Additional applications and benefits of such cells and methodology will be readily apparent to those of skill in the art after consideration of how the recombinant genes were isolated and employed in the construction of the compounds of the invention.

The FK-520 biosynthetic genes were isolated by the following procedure. Genomic DNA was isolated from *Streptomyces hygroscopicus* var. *ascomyceticus* (ATCC 14891) using the lysozyme/proteinase K protocol described in Genetic Manipulation of *Streptomyces* - A Laboratory Manual (Hopwood *et al.*, 1986). The average size of the DNA was estimated to be between 80 - 120 kb by electrophoresis on 0.3% agarose gels. A library was constructed in the SuperCosTM vector according to the manufacturer's instructions and with the reagents provided in the commercially available kit (Stratagene). Briefly, 100 µg of

10

15

20

25

30

genomic DNA was partially digested with 4 units of SanSA I for 20 min, in a reaction volume of 1 mil, and the fragments were dephosphorylated and ligated to SuperCos vector arms. The ligated DNA was packaged and used to infect log-stage XL1-BlueMR cells. A library of about 10,000 independent cosmid clones was obtained.

Based on recently published sequence from the FK-500 cluster (Motamedi and Shafice, 1998, Eur. J. Biochem. 256: 528), a probe for the JkbO gene was isolated from ATCC 14891 using PCR with degenerate primers. With this probe, a cosmid designated pKOS034-124 was isolated from the library. With probes made from the ends of cosmid pKOS034-124, an additional cosmid designated pKOS034-120 was isolated. These cosmids (pKOS034-124 and pKOS034-120) were shown to contain DNA inserts that overlap with one another. Initial sequence data from these two cosmids generated sequences similar to sequences from the FK-506 and rapamycin clusters, indicating that the inserts were from the FK-520 PKS gene cluster. Two EcoRI fragments were subcloned from cosmids pKOS034-124 and pKOS034-120. These subclones were used to prepare shotgun libraries by partial digestion with Sau3AI, gei purification of fragments between 1.5 kb and 3 kb in size, and ligation into the pLitmus28 vector (New England Biolabs). These libraries were sequenced using dye terminators on a Beckmann CEQ2000 capillary electrophoresis sequencer, according to the manufacturer's protocols.

To obtain cosmids containing sequence on the left and right sides of the sequenced region described above, a new cosmid library of ATCC 14891 DNA was prepared essentially as described above. This new library was screened with a new *fkbM* probe isolated using DNA from ATCC 14891. A probe representing the *fkbP* gene at the end of cosmid pKOS034-124 was also used. Several additional cosmids to the right of the previously sequenced region were identified. Cosmids pKOS065-C31 and pKOS065-C3 were identified and then mapped with restriction enzymes. Initial sequences from these cosmids were consistent with the expected organization of the cluster in this region. More extensive sequencing showed that both cosmids contained in addition to the desired sequences, other sequences not contiguous to the desired sequences on the host cell chromosomal DNA. Probing of additional cosmid libraries identified two additional cosmids. pKOS065-M27 and pKOS065-M21, that contained the desired sequences in a contiguous segment of chromosomal DNA. Cosmids pKOS034-124, pKOS034-120, pKOS065-M27, and pKOS065-M21 have been deposited with the American Type Culture Collection, Manassas, VA. USA. The complete nucleotide sequence of the coding

WO 00/20601

sequences of the genes that encode the proteins of the FK-520 PKS are shown below but can also be determined from the cosmids of the invention deposited with the ATCC using standard methodology.

Referring to Figures 1 and 3, the FK-520 PKS gene cluster is composed of four open reading frames designated *jkbB*, *fkbC*, *fkbA*, and *fkbP*. The *jkbB* open reading frame encodes the loading module and the first four extender modules of the PKS. The *jkbC* open reading frame encodes extender modules five and six of the PKS. The *jkbA* open reading frame encodes extender modules seven, eight, nine, and ten of the PKS. The *jkbP* open reading frame encodes the NRPS of the PKS. Each of these genes can be isolated from the cosmids of the invention described above. The DNA sequences of these genes are provided below preceded by the following table identifying the start and stop codons of the open reading frames of each gene and the modules and domains contained therein.

15 complement (412 - 1836) fkbW	
complement $(2020 - 3579)$ fkbV	
complement (3969 - 4496)	
complement (4595 - 5488)	
5601 - 6818 fkbE	
20 6808 - 8052 fkbF	
8156 - 8824 fkbG	
complement (9122 - 9883) <i>fkbH</i>	
complement (9894 - 10994) fkbI	
complement (10987 - 11247) fkb.J	
25 complement (11244 - 12092) <i>fkbK</i>	
complement (12113 - 13150) <i>fkbL</i>	
complement (13212 - 23988) <i>fkbC</i>	
complement (23992 - 46573) <i>fkbB</i>	
46754 - 47788	
30 47785 - 52272 <i>fkbP</i>	
52275 - 71465	
71462 - 72628	
72625 - 73407	
complement (73460 - 76202) [fkbN	
35 complement (76336 - 77080) <i>fkbQ</i>	
complement (77076 - 77535) <i>fkbS</i>	
complement (44974 - 46573) CoA ligase of loadir	-
complement (43777 - 44629) ER of loading doma	
complement (43144 - 43660) ACP of loading dom	
40 complement (41842 - 43093) KS of extender mod	ule 1 (KS1)
complement(40609 - 41842) AT1	
complement (39442 40609) DH1	
complement (38677 - 39307) KR1	
complement (38371 - 38581) ACP1	

```
KS2
     complement (37145 - 38296)
     complement (35749 - 37144)
                                       AT2
     complement (34606 - 35749)
                                       DH2 (inactive)
     complement (33$23 - 34480)
                                       KR2
     complement (33505 - 33715)
                                       ACP2
     complement (32185 - 33439)
                                       KS3
     complement (31018 - 32185)
                                       AT3
     complement (29869 - 31018)
                                       DH3 (inactive)
     complement (29092 - 29740)
                                       KR3
10
     complement (28750 - 28960)
                                       ACP3
     complement (27430 - 28684)
                                       KS4
     complement (26146 - 27430)
                                       AT4
     complement (24997 - 26146)
                                       DH4 (inactive)
     complement (24163 - 24373)
                                       ACP4
15
     complement (22653 - 23892)
                                       KS5
     complement (21420 - 22653)
                                       AT5
     complement (20241 - 21420)
                                       DH5
     complement (19464 - 20097)
                                       KR5
                                       ACP5
     complement (19116 - 19326)
20
                                       KS6
     complement (17820 - 19053)
                                       AT6
     complement (16587 - 17820)
     complement (15438 - 16587)
                                       DH6
                                       ER6
     complement (14517 - 15294)
                                       KR6
     complement (13761 - 14394)
                                       ACP6
25
     complement (13452 - 13662)
                                       KS7
     52362 - 53576
     53577 - 54716
                                       AT7
                                       DH7
     54717 - 55871
                                       ER7
     56019 - 56819
                                       KR7
30
     56943 - 57575
                                       ACP7
     57710 - 57920
                                       KS8
     57990 - 59243
                                       AT8
     59244 - 60398
                                      DH8 (inactive)
     60399 - 61412
                                      KR8
35
     61548 - 62180
                                       ACP8
     62328 - 62537
                                       KS9
     62598 - 63854
                                       AT9
     63855 - 65084
                                      DH9
     65085 - 66254
                                       ER9
40
     66399 - 67175
                                       KR9
     67299 - 67931
                                       ACP9
     68094 - 68303
                                       KS10
     68397 - 69653
                                       AT10
     69654 - 70985
                                       ACP10
45
     71064 - 71273
         1 SATOTOAGGO AUGAAGTOOT COAGGOGAGG 1500GAGGT3 GUGAACACOT OGCOGOTGOT
61 TGTACGGACO ACTIONGTOA JOGGOGATT3 OGGAACGAG3 TOATOGGGAA TARAGGGGGG
        121 TTACAAGATO CTCACATTGC GCGACCGCCA GCATACGCTG AGTTGCCTCA GAGGCAAACC
        181 GAAAGGGCC GGGCGTCCG CACCAGGGCG GAGTACGGGA CGAGAGTGGC GCACCCGCGC
50
```

			00000000000	201221115	2000000000	ACACCCTTCC	32777730703
	- 4 .			~~~~~~~~~	the second second second second		2202222222
	277.1				adriumizuded amaamataa		- 2002 00 2000 00 00 00 00 00 00 00 00 00
	752	TA03030A30	-3030AU = 914 a		7777758773	000000000000000000000000000000000000000	3,0%,003.0
_	• •	283223333	7770000A3A33A	10 0A 0 0 0 0 0 0	11.133774731		00%00000000
5	÷ - 1	27722233372	3330337337	1931337383		A33373373A	
	5.42	373A3A3737	ASCAMASSOT	3383773377	3333AA33T1	000A03A833	0070007070
	4.11	2023203200	T03AT00000	ASTABOOSTA	0000000000	2322327323	13828TA230
	331	3007738097	0364000033	30330A331A	3083080373	3838373007	CGATGGTGAT
		0A00000TT0		0007044030	GATGGGTTGG	Addacasas	3380300038
10	:		5000111100110	AGTOGGGATT	2000200000	200000000000000000000000000000000000000	333000AATA
1 57		3340035503				*********	GGTAGACGCG
	542	2007070000		UUUUAT LIAAA		77777777777	
	3.1	CTGCGTCAGA	TOCCAGTAGA	CCTCGTGGTG	GTACGGCCAC	202012121212	AGTICGGCCGG
	961	GRACOCCGCC	COGRECACO	CCTCGCGCGC	0736006307	3033330030	CTGCCGCGTA
	:00:	33733333713	TOGCGCAGGG	000000330A3	SAAGGTSAAG	ABSTTBBGAC	COTOCGCGCG
15	1081	CORCAGGGTG	COTTCCCAST	CONCRETE	07007X 7A00	TOSSSATSST	TOTTOCAGOTG
	1141	COAGOGOAGG	AGGTAGCCGC	CGTTGGACAT	ThreeTGACT	Websirgoson	CGAGCGGCCG
	1001	GTGGTAGCGC	TGGGCGACCG	ACGCGCGGGGC	10099970	ASCISSEIGA	SSCGGGTGTT
	1261		ACGGCGTTGT	CCGGCCGGGA		TAGAACGCGG	GGCCGGTGTT
				AGGCGTAACC	2030303AGC	ACCOMOTOGG	001000000
20	1321	GCCCTTGTCG	GTGGCGGCGT				CACCGITCCA
20	1391	GTCGTTGGCG	TACTOCTOSC	GGTTACCGGG	3973003370		
	1441	3000100030	ASCOGGATGA	CGAACTGGGGC	3.03.33.17	CACCCCTCCT	Tagrarragr
	1501	GGTGGAGGTG	TOGGGGAAGT	AGCOGTOGAT	CTGGATCCCC	333A070033	TGGGAGTGGC
	1161	CAGGOTTOTTO	SGCGTCAGCC	CTGCCCASTC	00000333703	973736006	TBBCCGCCGT
	1601	1000300370	STORGOTTE	COMGGCAGTO	BOTOTOTTGA	1313000000	CCGGGGACACG
25	1681	CAGCTGGGAC	AGACGGGTGC	ASTGACCSTO	0000007700	SBAGCAGGCC	3330037330
	2002		AGCAGGACGG	CGACTGCG3C	CASSSTGAGA	0000004660	castaceter
			CGTCCGACAC	CGAGGGGGAG	AMOCATOGAO	ACCOCCCACA	CGTGCGGATG
	1801	70703333300		GTCGCGCACA	GTGGNGAGGA		GOCOGCCATG
	1861		TGGAGGCTAG			0333400303	
- 0	1921	ACTGAGGCCC	CTCAGAGGTG		ATGACGGGGG		GGCGCTCCGG
30	1981	GGCGGTGCCC	GCGGCCGCCA		grocccgggT	CAGGGACAGG	TGTCGTTCGC
	2041	GACGGTGAAG	TAGCCGGTCG		CAAGGTGGTC	GTGACGAAGG	TGTTGTACAG
	2101	GCCCATGTTC	TGGCCGGAGC	CCTTGGCGTA	GGTGTAACCG	GCGCTCGTCG	TGGCGCGGCC
	2161	CGCCTGGACG	TGAGCGTAGT	TSCCGGCGGT	CCAGCAGACG	GCCGTGGCAC	CGGTCGTCTG
	2221		GCGCCCGAGA		CTTGCCGTCC	GCSTCCCGGG	CGGCGACCGC
35			GATGTGCCCS	CCCTCAGGCC	GGTGTCCGTG	TROGROGICS	TGGCGGACGT
20		GGTGATCTGG		GGTGGACGGC	GTASTCGGTG	GCGCCGTCGA	
				CGGTGGCGCC	337660360		GAGCGGGGCAG
	2401		ATGGTGGTGT			TSCSTGATCS	
	1461	CGARCCGGGG	TCGGAGGCGG	ATCCCCCCCCA	GCCGAAGAAC		
	2521	ACAGATCGAS		AGGCGGCGCC	GGTGCTGCCG	CACTOOTGTG	oncoggrade
4()	2881	GGGATCGACC	000	GCCCGATGCC	09392200033	TTCACCTCCA	CGGCCACCGA
	2641	moderaces	GCCAGGTACT	CCTCGTGCCG	GGTGGAGTTC	GGGCCGATCA	CCGAGGTACG
	chai	amanagaama	manasasasa	CGTGCACAGO	GGTCCACTGG	TOGOGCAACT	CGTCGGCGTT
	2.00	000000000	Laggaggaaa	Commandana	STECCHEATE	GCCACGCGCG	GCCACGGGCC
	2001	3000000000	0000000000	02003220003	00000000000	restactions	TCAGGTCGGT
45		CONSCINS	- recordance	* 000001.0000	25000000000	GCACAGCCGA	idageAdaee
45	_551	000000000000000000000000000000000000000	MIGUACAGGI	202101222	0.00100010	0.00000000	COCCCCTCAT
	1941	GGCGACGACC	6030033001	GGANGAGG . C	0.0000000000000000000000000000000000000	GCGAGCATCA	acm accccit
	3001	SGCACCGCCG	GCGGACAGCC	CGGTGATGTA	3313020133	SEGTECGES	
	3061	GACGSTOTGA	GCGGCCATOT	GCCGGATCGA	CGCGGCTTCG	CONTEGECCC	TGUGGTTGIU
	2101	GOTGOTOTGG	AACCAGTIGA	AGCACCTGTT	CGCGTTGTTC	GACGACGTGG	TCTCGGCGAA
50	53,83	CACGAGCAGG	AAGCCATAGC	GGTCCGCGAA	TGAGAGCAGG	CCGGAGTTGT	CGGCGTAGCC
50	5041	0m000000000	TGGGTGCAAC	CGTGCAGGGC	GAACACCACC	SCCGGCTCCG	CGGGCAGGGA
	2271	0.00000100	ma at comica	######################################	3000333770	GTGCCGAAGT	cagagagagaa
	3301		ACCEMENCE	70115119669	200000110	GCCGCGCGT	ggggggggaaa
	3301	GG. SAGGIDS	6661163168	omccoddci.	33307333000	1000000000	chaccements
	3421	cacasaacas	AGCAGGGCCG	CTCCGAGTAC		ACGGCCACGA	GWCGGGGTGWG
55	3481	CACCCCCCCC	CGTCCCGGAC	GCGACAACGA	CCCGACCGGC	GGCGAGGAGG	Manusususa
	3541	CAGCGGGGTG	AGGATTCCCC	GGAACGGCGG	CGGCTGCATG	GCGGCTCCCT	CGATGTCGTG
	3,533	GGGGGGACEC	GGAGGGCTCC	CTGACGTCGA	TOAGTGGGAG	CGCCCCGGTG	CCCGGCACCS
	1221	77.33337337	mcaacccsca	ACGGTATGGC	CCGGAGCACC	RORCCCCGCA	COGCGCGATG
	3331	######################################	CGGATTGTGT	CGCCTTGCGG	AATOTGATAC	CCGGACGCGA	CGAACGCCCC
60	i i i i i i i i i i i i i i i i i i i	10000000000	SOTAGGGGGG	CAMBGAGAGA	53.0m0450CG	GTCGGCCTTG	CCTGCCCTGG
60	: 51	Mullomenia	GOTAGOGGG.		20.033500	5155555119	

				30.03			
		A0000A00000	11371030003m	10000000000	3033300333	233TAT3303	GCCGAGGACG
		008300300	33333333333	03000744371	P CROTACGOCO	ACCOTAGCCA	GCGGGAGGGC
			100000000000000000000000000000000000000			000000000000000000000000000000000000000	GGGTTCCACC
-		7033337477	372777771		i Taballatia	27 27 22 20 20 20 2	
5	- "						100001000
	4141	2332733273	ACT000037			A STA SATGGT	
	4	AAGACCBBBT	77333543337	GA00003701	- 0%8003 %6 97	TESCORTORS	ATGCTGGGAG
	4261	ACOTOGOTOA	0001010000	BOTGACCAGS	ASTOCEASSE:	A03703A070	CACCAGCGGC
	4321	TT300000k33	mggT000030	CGAGTAGTGG	and the second second second second	2011000	SGTGTTGTGC
10	4060	370A33A300	TGABCCAGGA	GTTGTCGSTC	TOCAGGACCS	CACCCCCCCC	993376666
	4441	2828237232	COGTOSTOAA		TAGOGGCCCT	GCCAGCCCC	GACCACAGCG
	1511	3750363733	carcorggro		GTCGTCATGS	COCTOACCC	GGGAAGTCCC
	4561	2037203273	TGALATGOCS	AACCTTCACC		GTGCGGGGGA	TGAGCCCTGG
	4611	ACCGTACGTA	CTCGTAGAAC		0766060306	TESTCOTOCS	
15	468	CCACGCCGAC	007000000	G. LIGOGGST		DACGCCARG	
	47.1	033300° 'SA		GTGAGGGGGG		ACCGAGGCCG	GCGGCGACCA
	4501	GGGCCGCAG	CGTGCTCAGC	TOGGTGCTCT			AATOOGGOOG
	4561	CGGCGCACAG		ATCTGGCGCA			
	4921	CCTCATCGGC	CAGCTCCGCG	GTCCGCACCC			TGTCCGGGTG
20	4981	GGACGAGCAG		TOGTOCOGOA			TCCCCGGGCGG
_0	9202 8233	GTCGTGGGCT		AGGTCCAGCC			ACCACGGGGT
				TOGARGGTOS			TACCCGGCGA
	5161	0330030310			. 3.5533555	ACCOAGTGCC	
	5201	33AG3T03G3			AGTIGCAGGAA		
25	م بميد د	TGTCGGGGTC		GTGATGCGCT		GGAGACCTCA	
25	5281	GCAGGGCGTG	GGCGCGGAAG				TTCTGGTGCC
	5341	GGTCGNACAG		ACTOGTOGOT		GATGGCCCTG	
	5401	CCTGGGAGAT		TCCGCGGTGA			GCCAAGGCCG
	1461	TGAACCACTG			AGGGACTATA		ATGGTCCTGG
2.0	5521	CGAGGTTTCG	TCATTTCACA		GGCGGCCCAC		
30	5581		GAGGGACCCC				GAACGCCCCG
	5641	cccccccc			CTTTGGAGCA		GCTCCGTTCG
	5701	CCACCCCCCA			GTGTCATCAA		
	5761	GCGACCTCGC	CCGCGGCTAC			GTCCAGCCAC	
	5821	TGAACCGGGG	GAAGGAGAGC	GTCCAGCTCG	ATGTGCGCTC	GCCGGAGGGC	AACCGGCACC
35	5881	TGCACGCCTT	GGTGGACCGG	GCCGATGTCC	TGGTGCAGAA	TOTGGCACCC	GGCGCCGCGG
	5941	GCCGCCTGGC			GCGGAGCCAC		CACCTGCGGA
	6001	CATATCCGGC	TACGGCAGTA	CCGGCTGCTA	CCGCGGACCG		
	6061	RECETEGACON	AGCGGGGCTG	GTCTCCATCA	CCGGCACCCC	CONGRECCE	TCCAAGGTGG
	6101	GOCTOTOTAT	CGCGGACATC			CTCCGGCMTC	CTCACGGCCC
40	6181	TGCTGAAGCG		GGCCGGGGGCT		SGTCTCGATG	CTCGAAGCCC
	6241	1003102213			ACACGCGCTA	CGGCGGCACC	GCTCCGGCCC
			CAGCCACGCG			GTTCACCACG	CGCGACGGGC
	ését.	AGAGGAMONS	Temegggene	CAGAACGAGC	GGGAGTGGGC	TTCCTTCTGC	GGTGTCGTGC
	6421	"1,741,00000 "1,741,00000	CGGTGTCTGC	GACGACCCGC	GCTTTTCCGG	CAROGOCGAC	CGGGTGGCGC
45					AGGTGACGGG		
	6511	TOSTAGOGG	COTTONIO	00000000000000000000000000000000000000	CCTACGCACG	202020200	GTGCGGGAGT
	6.501	7001000000	COCCODAGE	COTTOS COOTTO	GACGCTGGGC	TTOSTTOSEC	AGCCCGGTCG
	6661		COCCOMMOIS	CCCCCCCTCL	CCTTCCACGG	031301100110	COCCOCCTOC
		31000001038	COXCOMOCOC	COCCCOGICA	AGTCCGTCCT	SINGONCOCO	20000000000
50					CCGAATGAAC		
20	25.45	nonsoducan	TOOCGAAGAG	CCCCCCCTCCC	GGGCTGAACA	TOROCOGRA	COLONICOTO
	0.247	0000000000	Tagracion.	CGGCGIACGG	OF CCCTTCC	1300001001	0.0000010010
	4901 47.51	accaccata	100 TC3055T	CCMCCMCAC TO	GACCGAACGC		0010000000
					GCCGTCACGT		
55					GTCGCGGTGC		
55	981	GGAGCCGTCC	CCIGGGTGCT	CTTCGGCCTG	GOGGCACTGC	LUTTGUGUGAC	MGGCGCGGCCC
					ATCAGCGTCG		
					GTGARCGGGG		
					7000000000		
2.5	7321	AGCGGCGGGC	TGCTCTTCGC	AGGCACCTTC	GCCTTCAACC	TGGCGGTCGC	CGCGGTGTCA
60	7381	TGGGTGGTTG	TOGGGCGCAG	GOGCCTCGAA	CCACATGACC	TGGACGAGGA	CACCGATCCC

	*****	7.039An3000)	200000000000	ngggggggg	0000AA0A00	TGATGACGCT	GACCCCATC
		3000000000	TGCTGGGAAC	CACCCCCCC	20202022	2233077007	GGCCTCACC
	25.61	7733333377		comemmence:	ng ladomora	AGGAGGGGAG	CAAGGAGATO
			000000000000000000000000000000000000000	2375722223	220213221		BOTOCAGGAG
5	16-1	1133322103	.33h2T233T	3333NA 3AT 5	AT1303333A	1 100000000	3013013600
		200022023	T0T30TA30T	9360337370	375753333	100000005.	0A00333AT0
	* 3 1 1	777277377	73273000307	370037.3003	TOTOTORAGE	necersons.	CBBBACGACC
	-121	333A73273A	T00000T000	3300323333	A003T00T03	ADSCOMSTOS	CITCTCCACC
	1901	/473073070	1337367332	0440907000	2862332723	33772333337	STACCAGGG
10	1981	7700707007	3363030036	SSTOTOOSCA	0700070000	1373030513	33033300TT0
	8041	3133133061	SECRODOSSE	GAGCOGGAAT	COCCTOCAC	7737775555	TOSTOTOTOS
	8101	2734237523	STOAKSTOCK	0379003333	3332A3TA23	007700AT37	23330AT630
	8161	TARTORGATA	ACCOTOTOCG	ACACGITICT	000111A0011	1332433731	CCCTGCGCGA
	34.1	1310313373	CIGAGCCGGC	TGCGCGCGCA	GACGGCCGAG	01000000000	GTGGCGTACT
15	3251	30000000	.CCGAGGAGG	GACAGTTIIT	0320720073	973033 77 5%	10360 30 606
	e 3 4 1	TOAGG.GULS	GAGATOGGGA	CGTACACCGG	CTACAGCAL	1107600736	CCCGCGGATT
	5431	33230000333	GGCCGTGTGG	TGACGTGCGA	TGTCATGCCG	ARGTEGECCE	AGGTGGGCGA
	8461	GCGGTACTGG	GAGGAGGCCG	GSSTTGCCGA	CCGGATCGAC	3000338003	GCGACGCCCG
	8521	SACOSTOCTS	ACCGGGGCTGC	TOGROGREGE	GGGCGCGGGG	COGGAGTOGT	TOGACATGGT
20	8581	STICATOGAC	GCCGACAAGG	CCGGCTACCC	CGCCTACTAC	3833333333	TOCCGCTGGT
	8641	AGGCCCCGG	GGGCTGATCG	TOGTOGACAA	CACGCTGTTC	7779300623	TGGCCGACGA
	6701	AGCGGTGCAG	GACCCGGACA	CGGTCGCGGT	ACGCGAACTC	AACGCGGCAC	TGCGCGACGA
	8761	CGACCGGGTG	GACCTGGCGA	TGCTGACGAC	GGCCGACGGC	STORCCCTSC	TGCGGAAACG
	ê E.C.L	GTGACCGGGG	CGATGTCGGC	GGCGGTCAGC	STCAGCGTCS	713333333333	COTOGOGGAG
25	3881	SBOTTOCASAT	GCAGGCGTTC	GACGCCGGCG	geggAAGegg	000000000000000000000000000000000000000	GGACACGCAG
	6941	GGGCAGTCGG	AGTOCGOGAA	GCCCGCGAAG	CSSTAGGCGA	COCCECE	GOGGTTGCGG
	9001	TOCOTACCO	GGAAGTCCGC	CACCAGGTGC	90000000000		STCCGTGAGC
	9061	CASTICAGGA	TOSTOGOACO	GGCACCGAAC	3ACACGACCC	300m33A361	GGCGAGCAGT
20	9121	TTCAGGTGCC	ACGTCGACGG	CTTCTTCTCC	AGCAGGATGA	TGCCGACGGC CGGGATCGGT	GCCGTGCGGG GAGCACGCGC
30	9181	COGAAGOGGT	CGCCCATGGT	GACGACGAGG TGCACGCCGG	ACCTCATGGG TCGCGTTCAT	erggerggre	CGCAGCGTCA
	9241	GCAGGTCGGC	GTCGGAGTAG	TCCTCCTCCC	CCGCGGGTGG	GATOGTCATG	GAGAGGTCGA
	9301	GTTCCTCGAC	GCGGCTGAGT GAAGTCCTCG	TOGGGACOGG		CCGGGGCCTGG	TCGCGCGCGA
	9361	GCGAGCGCAG	GTACATCAGG		GCGAGTCGAC	CGTGGACACC	GGCGGGCTGA
3.5	9421 9481	AACCCGCCTG ACTCCGGCAG	CGACAGGAGC	GTGGCCGCCT	GCTCGGCCGG	GTAGCACCGC	ACCTOGGGCA
ال ال	9541	GGTGGAACGC	CACCTCGGCA	CGCTCGGCGG	GCTGGTCGTC	GATGAACGCG	ATCGTGGTCG
	9601	GIGOGRAGIT	CAGCTCCGTG	GCGATCTCGC	GGACGGACTG	COMOTTOSSO	CCCCATCCGA
	9661	TGCGGGCCMG	CACGAAGTAC	TOCGCOACAC	CGAGGCGTTC	CROROGOTCO	CROSCGAGGT
	9721	0373370377	CTIGITICSCC	ACCGCCTGGA	GGATGCCGCG	STISTOGASC	GTGGTGATCA
40	9781	CCTCGCGGAT		AGGACCACCT	CGTCGTCCTC	CAGCACGGTG	CCCCGCCACA
10	9821	: SETEMMENT	CaggmoddAG	ACCAGACACT	TGACAATGGT	CATGGCTGTO	CTCTCAAGCC
	aaa:	SSSAGCGCCA	GCGCGTGCTG	GGCCAGCATO	ACCCGGCACA	TOTOGOTGOT	GCCCTCGATG
	aakt	:momoozma:	SCHTGGCGTC	GCGGTACGCC	CGTTCGACGA	CGTGTCCCTC	TCTCGCGCCT
	10001	googlogogl	SCACCMGMGC	GGCGGTCGCG	GCCCCGGCGG	CGGCTCGTTC	GGCGGCGACG
45	10081	maramaga or:	SSAMOSMOSO	GGGCACCATO	TOGGGCGAGO	COTOGTCCCA	GTGGTCGCTG
	3 0 1 4 3	COSES SESSE	Labadadagaa	CGCGATCTUC	TOCGCGGTCC	ACAGGTCGGC	GATGTGCCCG
	1,0001	COGLOGICON	3373355555	GAGCGCL133	O. MASTSCT	00003337400	GGCGTGGGCC
	10261	10000033733	TGCGGCAGGC	COGCAGGATO	COGACGCAGO	COCAGGCGAC	CGACTTGCGC
	10301	000000000000	GTGACGCCGC	GACCAGCATC	GGCAGTGACG	CGCCGGAGCC	GGCCAGGACC
5Û	10381	30300030000	GCACACGCAC	CTGGTCCAGG	TGCAGATCGG	CGTGGCCGGC	GGCGCGGCAG
	10441	COSSACGSCT	TOGGGACGCG	CTCGACGCGT	ACGCCGGGGG	TGTC33C503	CACGACCACC
	10561	ACCGCACCGG	AACCATCCTC	CTGGAGACCG	AAGACGACUA	001001000	02000000
	10561	GCAGTCGTCC	AGACCTTGTG	GCCGTCGACG	ACAGCGGTGT	CCCCCCTCCAG	COCCCCC
	10621	GTCCGCATCG	CCGACAGATC	GCTGCCCGCC	TGCCGCTCAC	TGAAGUCGAC	GUUUUUUGAG.
55	10581	TTCCCGCTGG	TCAGCTCCTT	CAGGAAGGTC	GCCCGCTGAC		00000000000000000000000000000000000000
	10741	ACGGTCCACG	CGGCCATGCC	CTGCGACGTC	ATGACACTGU	GUNGUGANUT	000000000000000000000000000000000000000
	10801	COGACGTGTG	CGGTGAACTC	GCCGTTCTCC	- 5666 LeVilla - 55666 5556	ulluludadau aaraaraara	- dandadadada
	10861	SCOSCOACTT	COGOGOAGAG	CMGGGGGGTGG			0.0000000
(1)	10921	AGTTOGCOGG TOACGOTCAG	ACGTGTCUCA	0106666666	REJUNDIEDOU		ACTOGACGGT
60	10981	TUACGCTCAG	BUJADULAGG	CCCGCHGCCG	Grouncone:	000,300,3100	

		and the second second second second		ga T 00033355		* ~~~~~	wammamaara
		an additional	Jugangu Jan		- Distantedit		TOTTOTODAG
		01A0A00A00	83.100M100				ACAGGTCGCG
		377283332	CASTOCSACC	1337077037	2003A33AA	00055005400	CGTGCGCGAC
		100000000000000000000000000000000000000	1108000000	0.0000000000000000000000000000000000000	AMPROTES IT		3AA3000003
5		VINE CARROLING CARROLING	30000000000	70000000000	ACCTTSSILE	3 JA 3 CA 3 3 7 7	ACAGGGGGGG
-		The second secon	2222222222		Accondaction	0.0000000000	GTTGTCGATG
				0.0000000	57033M030	. r. cantang	
		COCACCACCC	000000000000			COAGGCACCC	CGTCATGAGC
	:	3007008007	COTTOGREGUA	- Milad de la de la de		1003030030	GTCGTTGATC
		AT 70000753A	3083003307	000000000000	00000000000	CARCARCCON	SATOGGCTTG
10	* * * * *	77000078000	70305800A6	37.200033333	900900AT00	000000000000000000000000000000000000000	GGTCCGGGT
• '		COCCARTOR	00000000000		*******	COBARCOS	CGTGCCGAGC
				33 h 3T03300	AGTTOGTORA		2010000000
		naar na kaan	フェックははヒレアい			a da a a a a a a a a a a	CGACGTGTTC
	117/1	2.25.25.000	GSATACOSGG		SAGACCSTGG		CGCCTTCACC
		1233337707	CESCADUACC	CTCGATGAGG	gcggTGGCCG	TADOGATOGC	GGGCAGCGCG
15		BACGTGGCCG	70":0137%	AUJUGGGAIIG	3001033033	30000300AC	GRGTTGTGCC
	12331	370060A3TT	CGGTGGCGAT	0000000000	GOOGCOGTAA	SGATOTOOTO	GGACGTGTCG
		ACGAGTGTCA	CCGGGACGCC	GTGGCGCAGC	GCGAGCGTGG	TGATGCCGGT	SCCCATCACT
				CTGGTSGTTC	ACCOMMITTE	00000000000	GGTCACCATG
	12061	COCGCGCCGA	GCACGATCAG		ACCOUNTRACT	27.000.10000	
		GUAGOGAGTA	CGGGTCGAGG	ACGTCTTCCG		27.10300100	TTGCGGCCGA
20	12181	GGCCGAGTTC	GTCGGCGAAG	CCGAGCAGCA	CGTCGAACGC	38T3T33TC3	GCGAACGCGC
	12141	TECCCETCEA	GTOGAGGACG	CTCAGGCTGT	CCCGGTGGTC	73333 3333	TCCGGTGCCG
	10321	CGCACAGGGC	CGCCAGCCAC	GGGCCGAGCT	0303310033	DAGTIGOTIGG	TACTOGCCCT
	12361	0000000000	omaccoogga	TGGTCGACGC	AGATGAACGC	PROGREGAGE	AGGGTCTTCG
				maamaasasa	CGATGGCGTT	ur or maor co	TBCGGCAGCC
2.5	Y	3.543.4333.	CTTGCCGGC	TCG.C.SSSS			
25	11:3:	3000010000	GGGCAGCACC	GGCCCTTTGC	CCGAGGGCAC	CGAGGTGACG	GTGGACAGGA
	11541	0270033330	gacagagaaa	TOCGCCGGAT	CGGTCACCTT	GACCGGCAGT	CCGAGGAACG
	12601	CGATGCGGTC	CGCGAACGAC	GCCGCGTGTGGC	CGGGGTCGGT	GTCGCTGACC	AGGATOCGCT
	12661	CGATGGGCAG	GACCCTGCTG	AGCGCGTGCG	COTGGGTCAC	coccraraca	COOGCGCCGA
		TOAGOGTGAG	CGTGGCGCTG	TOGGACOGGG	CCAGCAGCCG	GCTCCCGACG	GCGGCGACCG
30	لمث سد			ATCACGCCTG	CGTCGGCGAG	GGCGGTCAGA	STGCCGCTGT
30	12781	CGCCGGTCCG			TCGTCGGCAG	CCGGAAGCGC	GGATAGTTGT
	12841	CSTCCTCGAG	GCGCGACATC				
	12901	GCGGACTGTA	CGAAACCGTC	TTCATGGTCA	CGCCGACACC	GGGGACCCGG	TACGGCATGA
	12961	ACTOGATGAC	GCCGGGAATG	TOGCCGCCGC	GGACGAATCC	GGTACGCGGC	GGCGCCTCGG
	13020	COECTOAA	GCGGCCGAGC	GCGGCGAACC	CGTCGTGCAG	CTOGGTGATC	AGCCGGTCCA
35	13081	TOATCACCTC	GCGGCCGATC	ACGGAGAGAA	TCCGCTTGAT	STOATGTTGG	CGCAGGACCC
	13141	TGGTCTGCAT	GTGTCACCTC	CCTTTCGTGG	COGGAGOTGT	CITGGTGGTG	CCGCTCGGGG
				GCTCCCTGTC		ARRATOTOGT	CCGCGGTCGC
	13201	CGGCTTCCST	TCTCATCGCA				AGOGGTTGAG
	13261	STCCGCGGAC	AGCACGCCGG	CCGGCGTGGT	CGGGCGGGTC	TOCOGOGGG	
	13301	CAGGGGGTCC	AGCCGGGTTC	CGATCGCGTC	CGCCTGGCGG	GCGCCCGGGT	CGACACCGGC
40	13381	AACGAGTGCT	TOCAGCCGGT	CGASCTGCGC	GAGCACCACG	GTCACCGGGT	CGTCCGGGGGA
	13441	CAGCAGTTCA	COGATGOGGT	CGGCGAGTGC	GCGCGGCGAC	GGGTAGTCGA	AGACGAGCGT
	19891	GGCGGACAGT		TOGOCTOGTT	GAGGCCGTTG	CGCAGCTGCA	CCGCGATGAG
		CGAGTCCACA	000.10.10000	CONTROCCO			namessection
		- Lambiloumen		TOTGCCGGAC	010010000		CCCCTTCCTG
	13621	GCCCAGGACG	300301135551	101000000	GMGGGGGMG1	401.0001.00	GGGGGTTCCTC
45	13681	CTCGTTGCGG	GOSCTCCGGC	GGGCCGACGG	CT Tabactaa	Number and the	GUGGGAGGIC
	13741	DGGCGGCA73	TOGCCCCCCA	CGGCGACGAC	Annacacani	CCGGTGTGGA	CGGCGGGGGTC
	13211	0720276060	ATGCCCTGTT	CGGCGGTGAG	CGCGCTCGCC	CCACCCTTGC	GCATACGGCG
	15661	concentages	COSGTCAGGT	CCGCGGTCAG	GCCACTCGCC	TGGTCCCACA	GCCCCCACGC
	13921	GATOGAGAGG	commandance	OwnCadus as	adagamentes.	GCGAGCGCGT	CGAGGAACGC
50		SAL SSASAGE	000000000	CCTGACCGGG	campoonida	1010233023	occicciona
50	13981		JUSTAGITA	0010800033	00.00000000	MONTECOGGGG	1000000000000
		GAIGAIGAAT	GCGGCGVGCT,	(della lana	361343553	130M36.666	A-3000000000
	14101	GGCCTTGGGT	TTGAGGACGG	TGTCGATGCG	GTCGGGGGGTG	AGGT TGTCGA	GUAGGGGGG
	14161	300020000	neggegggggg	GGAAGACCCC	GGTGAGGGGT	TGAGGGATGT	GGGCGAGGGT
	14231	norgadalor	magmagaaam	CGCCGACGTC	GCAGGGGAGG	TGGGTGCCGG	GGGTGGTGTC
55	13001	GGGGGGTGGG	.55155555	COLCOLLOG	GTGGGGGTGG	TTCAGGTGGC	GGGCGAGGAT
رز	14251	SCCGGCGAGG	GIGCGGGAGA	GOOGGE TAGGE	0100000103	CCCTCCCCCT	CCAGCGGCCG
	14341	GCCGGCGAGG	GTGCCGGAGC	C6C661671	OMCOACCACC	CCC1CGGG1	CCAGCGGGGG
	14401	CGGGGACCGTG	AGGACGATOT	TGCCGGTGTG	UTUGUUGUGG	UTUATEGICS	JUMB COULTS
	14461	GOGGACCTGC	CGCATGTCGT	GCACCGTCAC	CGGCAGCGGG	TGCAGCACAC	CGCGCGCGAA
	14501	c≗ggccgigc	AGCTCCGCGA	TGATCTCCTT	GAGCCGGTCG	GGCCCCGCGT	CCATCAGGTC
60	1,1521	GAACGGTCGC	TEGACGGCGT	GCCGGATGTG	COTOTTCCCC	ATCTCGATGA	ACCGGCCACC
UU	74707	05230601000		300000000000000000000000000000000000000			

	14641	000000000000000000000000000000000000000	A00003A000	ACCOCTOGAD	3837778003	STGRECERET	TGAGGAGGAG
	14001	00003800000	GGGAACGCC	23323FAC30	0373073033	GAATOGGOOM	3200000000
	14761	3T00A33T20	ACCACATOGO	3077030333		COGTACACCT	
5		000000000000	ATOTOCOGG		i dibebera a a a a di Tanana a a a a a a a a a	in radio salada. Persanan	30A10A00A0
-,	[455]		33303314300	0330343371		RADDAC BCGG	
	-4241	9377777803	GROSCOSCOT	303 3 3144037		190%7700330	CGAGCATCCG
	18000	GTGGTC5600	AMBACCOMES	GGCCGAAGCC	397300A73	AGGCCGAAGA	0003670600
	15081	caaractwak	CCGGAGAGGT	036030033	CTCCAGGACG	ATROCCCCCC	0070900300
	12151	9A93A09000		AGGTGGGGAR		ACATOGOGGA	2007022000
10		000000000000			0330303A33	33333030336	301000000A
1.77			AUMILISHIOD		000000000000		
			Giranaggar.				GOGACGTGTC
	15351		ACGGTGAGCG	GCTCCGGCAC	00000000000	0.13000000	0384003300
	13361	3000000130	COCHORCOCS	GCTCGCCGAG	TGCGACGGCC	ATROCOCOCT	3770533330
	15421	32.3337323	COGGRETCGG	TOTOGROOMS	GACGAACCGG	200000000	0330016360
15	18461	3673767763		cogogogogom	3603A36000	3033736737	SCACGAGCAG
-	10.5			GGGCGGTCAG		STSASCGCAC	GCGTCTCGGC
					CGTGATGACG		
	15601	0.5000000000	- AUGULATUAG	Cooundacin	COLORIGADO	1000000000	TOGOGGGGAC
					GAGACGCATC		CGACGGGTGG
	15721		CGGGTGCGGA			AGTIGGCCGG	CGGAGTCGGT
20	15781	SACGCGCAGA	CTCAGCTCGT	CGCCGTCACG		GOTOGGAGOA	TGGCCGAGCC
	18841		AACCGGGGCCC		GAACGGCAGA	COCCAGOSO	TGTCGTCCGG
	15901		GCGACGGCGT				0500354500
	15961	270010F.30	2001.003001	0000000000	CGCCACCTCG		TOTOMODATO
						materia and	CATCCUSCAG
~ -			GCCCGCAACC				
25	16081		AACCCCGAGA		cacagegata		ACTGCGAGAA
	16141	COCCTOCACA	COGACAACAC	CGGGGGGTGTC	SSSGCTGTCS		TGCCGCTGGC
	16201	37300333370	CAGCTGCCCG	TGCCCTCGGT	ACGCGCGTGG	ACGGTCACCG	SCCGCCGTCC
	16161	ggografica	GCCCCTTCCA	CGGTCACCGA	CACATOCACO	GOTGOGGTOA	COGGCACCAC
	16321	1100000000	modimoidor	COMOGRACEE	TETTCCCCCAR	CCGGTCTCGT	dadogacdog
30	16361	2182223334	TCCACAAACG	0010010000	ninnanginn	STECCOCCE	COCCOMCEMO
50		DARGUAGE A	100ACAAACC	COSTRCTOR	CHOCAGOLICO		CACCATCGTC
	16441	AGCCAGCCAG	Geblemeloc	GUARIURUR.	COGGCCAGTG	ADMACAMON.	
	16501	GGCGGGGCAGC	GCTGTGACAG	CGGCCAGCAT	CacAroloct	GCACCCGTCA	Miccipactes
	16561	CGACAGATCG	GTGGCACCGG	CCGCCTCCAG	COAGTACCGO	CTGTGCTCGA	ACGCGTACGT
	16621	GGGCAGATCC	AGCAGCCGTC	COGGCACCGG	TTCGACCACC	GTGTCCCAGT	CCACTGCCGT
35	16681	GDCCAGGGTC	CACGCCTGCG	CCAACGCCGT	CAGCCACCGC	TOCOAGCOGO	CGTCACCGGT
	16741	CCGCEACEAC	GCCACCGTGT	GAGCCTGCTC	CATCGCCGGC	ASCAGCACCG	GATGGGGACT
	16801	0000000000000	A A CACCCAACC	CATCCAGCTC	caccaccacc	SCSTCCAACG	CCACCGGACS
	70001	TOGOTOL MMG	AACACCCACC	NOGOCOCCE.	CACCGGCTCC	orologolgo	CGCTCTCCAC
	70007	MUSUMUMILU	COSIMDOMOI	ACCOCTONIC	2000000100	0.00000	COTTGGCCAG
4.0	16901	BETUGACCAD	CACGCCACUS	المناد واوال والاس	CCCTGCCACC	DECIDENSIN	
4()	16981	TTCATCCTCG	ATGGCTTCCA	CGTGGGGGGGT		TAGTOGACOG	CANTURGUE
	17041	CACCCGCACG	CCTTCGGCCT	CATACCGCGC	CACCACCTCC	TCCACCGCCG	ACGGGTCCCC
	17101	CGCCACCACC	GTCGAAGCCG	GGCCGTTACG	CGCCGCGATC	CACACACCCT	CGACCAGACC
	17161	GACCTCACCG	GCCGGCAACG	CCACCGAAGO	CATCGCTCCC	CGCCCGGCCA	GTCGCGCCGC
	17001	Campagagas	CTCCCCLITC	CCACCACGCS	geoggaggmag	TOGREGOTER	GGGCTCCGGC
45	1.001	CARCONCCION	222222777	accommoder.	amananatmo	ACCGCGTCCG	GOLOGIOCO
40			GUUGUGAIJI	CGCCCTGGGA	0.0.00000000	At COMCOCC	OCTICON COTO
	1 341	ATGUSCOTSC	CACAGUGUGG	CCAGGGIONE		CAGCTGGCCG	GC.GGACCAC
	17401	CTCCACCCCC	TOCGCCACAT	CCGGCCGCGC	CHACHICICC	COCACATOOC	AGCCCGTGTG
	17461	CGGCAGCAAC	GCCTGAGCGC	ACTOCTOCAT	ACGCGCGGCG	AACACCGCGG	AGTGGGCCAT
	17521	GAGTICCACG	CCCATGCCGA	CCCACTGGGC	GCCCTGGCCG	GGGAAGACGA	ACACCGTACG
50	17681	CARCTROTO	Andgedagen	COSTCACCOS	GGCATCGCCC	ABCAGCACCG	CACGGTGACC
2.0		0000100100	22200000000	CONTROCCOS	1303201616	GCCACATCCA	CACCACCOCC
	- Ca	GARGAGAGA	COCTOCCOM	CCMMCCCCCC	30000000	CTCACCTCAC	or conscoor
	1 1	a calemant at	COCHOCAGOC	UCTOURDUIG	000000000000000000000000000000000000000	ordano.com	CACCACCAC
	- 61	CACCGGCAAC	GGCACCAAJC	CGTUARCARC	Complete	CGCGACGGCC	UAGGAACAGG
	17821	CTCAAGGATC	ACGTGCGCGT	TOGTACOGCT	CACCCCGAAC	GACGACACAC	CCGCATGUGG
55	17881	COTABOOCET	GACTCGGGGCC	ACGGCCTCGC	CTCGGTGAGC	AGCTCCACCG	CACCGGCCGA
	17941	CCAGTCCACA	TGCGACGACG	GCTCGTCCAC	ATGCAGCGTC	TTCGGCGCGA	TOCCGTACCG
	18001	-07@53004@5	ZODECCEDE	TURULUCGO	GACACCOGGC	GCCGCCTGCG	CATGACCGAT
	1 0 0 0 0 1 2 0 0 0 0 0	-00010#10	- NOON - D- 16M	7 21.0000000		TOOTGOOGGT	Angmagaata
	70.01 TOUDT	GITUGHCITU	AUGUARROUMA	SOMEWHOLD CO.	-2740015W000	CACCCAACCC	CCTCCACCAC
20	18121	AATGGCCTGC	GCCTCGATGG	GATUGUUCAG	0010010000	GTCCCGTGCG	ON OCCUPANT
60	13181	GTOCACATOS	GCGGGGGGGA	GTCCGGCGTT	CACCAACGCC	TGCTGGATGA	CACGULGUTG

	11.41	P3A 003 3 3 1 1 2	77333333333	ACACCOCCT	0343334333	TOOTSATTOA	CCGCCGACCC
	14414	2220520000	CCCAAGAGAGG	TOTGTGTGGTT	30007733330	T233A3A600	GOTOCAGOAO
	18371		2222222222	00071000000	3003773300	GOGTGGGGGA	ACCCCCCC
			-3593835370				* *******
5				. Adda a a a a a a a a a a a a a a a a a			
-,	1:4:1	CATGAGGGT	Muricus access		-03/43.50.00	ALCOCCA.	GTGCGTGCCC
	18541	\$300T33T30	AggguGAuum		Managara a	alum	ACGCCGGTCC
	1=611	2703833228	TROAKGTAGG	Mawa nasan	22.20.00000	773030730A	maddaAmegA
	15861	300374.0003	TODASGTOOG	CGCCGACGCC	GTACCCCTAC	3A3AA33CGC	COATGAAGAG
	1 7 1 1	222222222	7700000000	-3 73 7303333	-0A03AT3300	200000005A	ADDOCTOOOA
10			ABBABBATBB	GOTGOTGGGG	3730 A73 323	737333 373 A3	GGGGGGTGAT
	18841		GOGGGATOGA	AGC0330330	37033A3A33	AAGCCCCCC	GSTCCGTGTC
	1 - 4 1 1	0.1280000000	373%330030	A00007000A	3003033703	3310333 AAG C	COSTGACCGC
	i e aci	3700000000 370000000000	amamaaraar	TGGGGGAGAG	GT03T03330	SASSTSACGO	pageggggag
		2000000000000			mm 10001000		
1.5			MI SULUMUSM	TOBUUMBUBB	TELLOLORISTO TAMBOL ALLES		0737999AA0
15	19091	madhina ada i	30330. Junu	CGACCAGAGC	and the second and the second second second	ou conscion	7930000003
	19241	13717333743	TOGRAGATÃA	6097030936	CASTOSSACA		53993 m 3103
	12.00	7773733737	TOGACGGCGG	TCAGCGAGTC	GATACCCAGT	100703AAG3	JUGGCGT COGC
	19061	3353835777	400000000000	CGTGGCCGAG	-CRCCGCCGCC	30000030	GGACCAGTGC
	19321	020000000	STSTSSSSST	CAGCGCCGGA	CATGGTGCCG	ABROSSTOGG	CGAGCGGAAC
20	19351	3330079333	3003003330	GCGATACGGC	GOGGCGCRGA	TOOOOGAAAA	GOGGCGATGT
~~	19441	37203333373	ASSTCCATOS	TGGCCGCCAC	3303AA0303	900000000000	gaaceaasac
			00000000000	CACCGGGGGGA	0.50000000		33036ACAC3
	19531	TTOOMSCAGS	noodinideen			74474000000000000000000000000000000000	
	19861	3370033773	GTGCCCCTCA	1907601661	- Undiction of		
	19621	a a comprame	AGCGCGGGUA	6.001.1331	AlgadadAac	GTCGCGAGTO	COTOGREGAR
25	19681	0000000000	GCCGAGTAGT	TGCCCTGGCC	3033333333	ADGATGCCCG	CGACGGACGA
	19741	GTAGAGGACG	AACGAGCGCA	GGTCCGCGTGTC	CCGGGTCAGC	TOSTICAGGT	GCCAGGCGCC
	19301	GTCGGCTTTG	GGGCGCAGTG	TGGTGGCGAG	cccctccasa	GT GAGT GCCG	TGGTCACGCC
	19561	enderesade	ACGGCTGCCG	TGTGGAAGAC	CCCCCTCAGC	0010T300GG	CGGCGGGGGAG
	1992:	2220323333	AGCTGGTCCC	GGTCGGCGAC	GTORCAGOGG	ATSTSSACAC	CGGGAGTGTC
30	19981	00000000000	TOGOTGOGGG	ACAGCAACAG	GAGGTGGCGG	GCGCCATGCT	CGGCGACGAG
	20041	ATGCTGGGGG	AGGAGACCTG	CCAGCACACC	CGAGCCGCCG	GTGATGACCA	COGTGCCGTC
	20131	CGGGTCGAGC	AGCGGTTCGG	GCGTTTCCGC	GGC350037G	COSCUSANCO	GCGGCGCTTC
	20161	GTACCGGCCG	TOGGTGACGC	GGACGTACGG	CTOGGCCAGT	ET CGTGGCGG	OGGCCAGCGC
	20221	00000000000		CGGTCTCCAC	CAGCACGAAC	COSCOCGGGT	GCTCGGCCTG
25			3001000100	CGACCGCTCC	TCCGACCGGT	COCCOTOGA	TCCGGACGAC
35	20061	33003110030	ACGAGGCCGG				CGAGCACGAA
	20341	anadalesi.	TOOGCAGGGC	CGTCCTCGGC	GATCACCOGG ATCCGCGCGCC		GCGCGGAGAC
	20401	STOGGTGASC	2337A03701	CGTCGAGGAC		GBTTTCGGGGA	
	20451	CARCETETAC	GCGTCCGCAG	GACCGGGCCC	GGGAGTGGGC	AGCTCCGTCC	AGGAGAGGCC
	00521	GTACAAGGAG	TTOCGTACGA	DGGCGGGGTC	GCCGTCGACG	TTOROGGTC	GCGCGGTCAG
4()	00581	0300300A03	STORCORCOG	GTTGGCCGAC	CGGGTTCGTC	GCATGCACGG	CAGCGCCGTC
	25641	COGGCCCCTGA	GTGATCGTGA	CGCGCAGCGT	GGTGGCCCCG		ACCGCACGCC
	20701	GOTCOMOGRG	AACGGCAGCC	GCACCTCCGC	TTCCTGTTCC	GCGAGCAGCG	GCAGGCAGGT
	20251	chocrocite	accaramesa	ACEGOGOOGS	-GTGGAC 10CA	TRGTGCGGCG	TGTCGTCCGC
	20221	andmoraced.	GOGETOMODE	COTOGGCGTA	CAGGGTTTCG	COGTOGOGO	AGGCGGTGCG
45	20022	07.000000000000000000000000000000000000	11000000000	CGTAGCTGTA	GOOGGTOTOG	GCCAGCGGCT	CGTAGAACGC
-15	20001	-0707000130	* CCCCTCCCC	canceagess	oggoowogog	GGCGGCGGGA	COGCOGCGAC
	200941	3010Mc3103	2000010000	TOCOCOTAGE	STRONGSTOR	CASCTGTCCG	TGCCCTCGGT
	21002	2011000000	- 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-000001000	200000000000000000000000000000000000000	GCCCCTTCGA	CGSTCLCCGA
		MUDDO La La a	ALGO LUAL L	3000000111	00000.00100	***************	CENT CT TOODS
	11111	CACATTCAACC	3030063 /cm	- COBBONCON	0100000000	TOGREGACCA	0007700000
50	11131	CACCCCCCAA	CCGGTCTCGI	CACCESCUES	OMIGMULMSC	TOCACAAACG	CCCTACCCCC
	31241	CAGCAGAACC	STGCCCCGCA	CCGCGTGATU	Restances	SGATGCGTAC	GCAROGRANI
	21301	CCSSCCAGTG	AGAACAACAC	CACCACOGTO	STOGGGGGGG	ROTOCTOTGA	Cabbabbab
	11361	CATCGGATGC	GCCGCCCCGG	TCAGCCCGGC	CGCGGACAGA	TOSSTOGCAC	Cadececcana
	55 105		according and	ATTORNESSO.	-GGTGGGGCAGA	TIGAGIAGOO	GTCCCGGCAU
55	11121	23377723122	LOCGTATACCO	AGTCCACTGC	CGTGCCCAGG	GTOCACGCCT	GCGCUAACGU
	93.5.13	20000000000	-AGOMADORAGO	CGCCGTCACC	GGTCCSCAAC	GRUGUURUUG	1010200010
	21.561		-naranatana	COGGATGGGG	garacaaraa	ACGARCACGG	ACCOSTOCAG
				angogangga	GCGACGCAGG	TTCCGGTACC	AGTASCCCTC
	2.00.		200000120A		CACCGTGGA	CACCAGGCCA	CCGACCCGGT
10	21/21	ALCUMUCUSII		MOGCOCIO	CF F CM 2 CA C	TOGATGGCTT	COACGTGGGG
60	21781	COCGCCGGAA	ATCCCCTCCA	GIMCCICSSS	CWW0.03100	1007.00041	50,,00,1000

	+ :		****	00000000000000000000000000000000000000	0.3001.00000	100000000	00000000000
	3300	- J.J.J.JJUNJ			0000001001	-5000011233	SUTTURIACTO
		- 000000000		000000000000000000000000000000000000000	00000000000	ACAGTOGRAG	ACGGGCCGTT
	11361	addidddddd	W. POWONCE	- USTUBAUDA:			ACGCCACCGA
-		AGUCATOSON	00000000000	208020000	Sadahi'da	TRROTTROACA	AGGCCACCAC
5			T00T0AA0G0	- Kiomoossaas	3300NCA0A1	000000000A	1010300013
		33 A37 37003	ACCACCGCGT	20290A03A0	000AT30300	T300A0A303	CGGCCAGGCT
		CACCGCGACC	GCCCAGCTGG	COCCCTGGAC	JANOTOCAC:	CGCTCCGCCA	CATCCGGCCG
	22261	DGCCAACATC	TOCCGCACAT	CCCAGGGGGG	GTGCGGCAAC		CACACTICTE
	22321	CATACGAGCC	GCGRACACCG	CAGARCACGC		AGACCCATGC	CCACCCACTG
1 ()	223331		COGGGAAAGA	CGAACACCGT	ACCCCCCCT	TOCACCOCCA	CACCCATCAC
	1041	000000AT00	CCCARCAACA	COGCACGGTG	ACCGAAGACA	GDAGGGTGAG	GCACCAACCC
	12891	27333333333	SOGGCCACAT	CCACACACC	COOCGOGGASA	TACCCCTCCA	GOOGOTOCAO
	22561	craccoccasa	AGACTCACCT	CACTOOGAGO	CGACACCGGC		ACCCATCGAC
		AGCCGACTCC	CCACGCGACG	GOCCSGGAAC		ATCACGTGCG	CGTTCGTACC
15	02881		AAAG13GAGA			2222237266	GCCACGCCCG
	22.41		AGRAGITOCA		GGTCCAGTC:		AGGGGTGGTG
	22831	DACK JOAGO	GTGTTGGGGG		COSCATORCO		TGATGACACC
	22861			GCGCATGACC		TTCAACGAAC	CCAGCAGCAG
	22921	-0000M0.0000	COCTCCTCCC	CGTACGTCGC	CAGNATOGOS		
20	22931			GCGCCTCCAC			
~0	23041			TGACGCGCTC		COGTTGGGTG	
				TGACGGGGGA			
	23161	G:16CGC1CG	ADADDOLDON	GCTTTTCGAC		TOOGGGGGGA	
25	23221	9916000100	GCCGCGTCAG	CGAACGCCTT	- academaran		
23				TCTGTGGTGA		GTGACACCAC	
				GCAGCGCCTG			
		CGAACACGCC		TGACCGCCGG			
	23461	TCCGGCGAGC	ACCGCGGGGCT	GTGTGCTGTA	GGGGGGGAA.	CCGCCCAGGT	
3.6	23521			CGCCGACGAA			
30	23591	CGGCACGATG	CCGGCGTGTT	CGAGCGCCTC	CCAGGCGATT	TCGAGGAGGA	
		CGGGTCGAGT	GCGGTGGCCT	CGCGCGGACT	GATGCCGAAG	AACGCGGCAT	CGAAGTCGGC
	23701			CCCGCCCGGT			
				TGGGGAAGTC		TOSOGGCCGT	
	23821	CTGCCACAGC	TOTTCCGGTG	AGGTGACGCC	GCCCGGCAGT	CGGCAGGCCA	
35	23881			CGGCGCGCAG			
	23941	GTCCTTGACC	GACGTCCGCA	GCGCCTCGAT	CASSTCSTTC	TOGGCCATCG	
	24001	TCAGCACGTG	CGCGATGAGC	GCGTCTGCGT	CCATGTCGTC	GAACAGTTCG	
	24661			GGTGCCTGTG		ACCGCCGTCC	GGGGTCCCGT
	24121	TSTCGTCCGG	GGTCCCGTTG	ACGTCCGGGG	CCAGGAGGGT	CAGCAGATGA	
40	24181	CGCCGGCGGC	GGGATAGTOG	AAGACSASCS	TGGCCGGCAG	OGGAATGOOG	AGGGCCTCGG
	24241	AGAGCCGGTT	GCGCAGGCCG	AGCGCGGTGA	GCGAGTCGAC	CCCGAGGTCC	TTGAACGCCG
	24301	TGGTGGCCGT	GACCGCCGCC	GCGTCGGTGT	GGCCCAGCAG	GGGGGGG	GTGTCGCGGA
	24361	CGACGCCGAG	CAGCACCTGT	TCCCGTTCCT	TGTGGGGCAG	STCCGGCAGG	CGTTCCAGCA
	24421	GGGAGCCGCC	GTCGGTCGCG	GAGCGCCGGG	TGGGGCGCTG	GATCGGTCGC	CACAGEGGTG
45	24481	ACGGGTCGCC	GGGCCCGGGT	GGGGCGGTCS	CCACGACCAC	GGCTTCCCCG	GTGGCGCACG
	24541	CGGCGTCGAG	GAGGTCGGTC	AGCCGGTCCG	CCGCGGCGGT	GAACGCCACG	GCCGGCAGGC
	24601	CTTGTGCCC	SCGCAGGTOG	GCCAGGGCCT	GGAGCGGTCC	GGCGCCTCG	CCGGACGGAA
	24661	CGGCGAGAAC	GAACGCGGTC	AGGTCGAGGT	CGCGGGTCAG	GOGGTGCAGT	TOCCAGGCCG
	2,223	ACTOGGOGGT	GOOGTOOGOG	TGGACGACCG	CGGTCACCGG	GGTTTCCGGC	ACTGTGCCCG
50	31781	3070672000	SATCACTTCS	GCGCCGTGTC	CGCCGAGGTG	TOCGGCGAGT	TOCTOCGAAC
2.0	23831	COCCOCCO	GLOCAPOSTO	TCGCCGTACG	AGGCCGCGGC	CGTGGTGGGC	GCGGCGGGGA
	24901	2222222222	CACTTACLAS	CGCCCGTCGG	20202020	STACGSTTCS	TOGAGGGGG
	03921	1010000000	70000000000	GGGGTGACCG	TGTCGGTGGT	3000203160	ACGARCOGRO
				AGTGCGGCGA			
55	120021	750075000	30101C0M3C	GCGGTCCTCG	000000000000000000000000000000000000000	000000000 000000000	TOUROUTURE
U U	20051	MUNUUMUUMU	0000000000	ATGACGTCGG	GGIGBIGENS	- 000001M000	WCCCLDargos
	45141	BAUUGGATAC	COGGREGATION	#1G#C3:C33	0001000010	Samach dama	ACCORDED TO THE TOTAL TO
	45301	330033CCCCT	961000066T	GCCGCCGGGG	DICEOMCOC.	Paramaraa Paramaa Paramaa Paramaa Paramaa Paramaa Paramaa Paramaa Paramaa Paramaa Paramaa Paramaa Paramaa Paramaa Paramaa Parama	COCCOURAGE
	20061	GUUGUACUTC	COCOTOCOGG	COCGTCGTGG		0010410400	ACCCCCATCC
60	20321	GAGCCACCGG	JUGTOUCAGT	TCGTCGGCGA	GCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CM CCC CA CCC	7.000001000
60	25381	CGTGGACGAA	DUDDADADTSS	AGTTTCGTGG	JUCCUU 100.	010022260	ACCCCG010M

			0/4/2007/800	~~~~~~~~~~		0000870000	000000TTG0N
		and the fact and an engineering	- Programme and a second	-003723A373		3303700073	202102222
		マース・マンチャ はんしん					
		is Densida	GROTTOGGOG	ondhodd.g.		1040000000	
~		1254 7775	الجامع للأصالة المؤتمون والوائر				300303A037
5	21:51	1320133773	00033730703	9909902A33	30000233232	33202022.	TCBSYBGTGG
	2.5	7987960990	SAAGCCGGAG	30 3T 332933	T00MT0T003	-3703003 7 00	3700363567
	15000	GGACGCGCAC	GGCACGGCGT	CCGGTGTCGT	033303333	GACGGTCACG	CGCACCTGGA
	LEBRIT	0330000037	SGCSSSCAGG	ACCAGOGGTG	TOTOGREGRE	CAGTTCCTCS	AGCAGGTCGC
	. 3 3 2 1	A3000000000	3703303000	CSTCC3GCCA	ATTOCABGAA	3333337003	SSCACCACTA
10	11891	2222222222		CCGGCCAGCC	ATGGGTGGGT	~ ~ ~ ~ ~ ~ ~ ~ ~	
1 (GACGGAGTGA	T0GG33A303	-00A000A000	-0000000000000000000000000000000000000	
		. Amachache	77706106646	GAA600T003	- 13000300001	- SETETTGATE	DOGTESTOSA
			TOCGAGGCCG				CASTASOGOT
	14171	M.T 00700AA	GGCGTATGTG		3733337332	2772303333	ADGAGEGEE
, -		00010703A0	3330A03003		000000000		ABBCBBTGGA
15		0000000000	GOGGCGAGC	GTGGCGAT		JAMUJOGGGT	AUDAGCACGG
	2.6042	337333333T	SOMODTCOMC	AACACGGTGT	OAC003332I0		373A333CC.
	3.641.1	T0000AA600	TACGGGGTGG	CGCATGTTGC	GGARODAGTA	0000000000	ABCBBBBBB
	16461	0387007600	TTCGTCGGCG	GTGGAGAACC	ACGGGATOTT	3770000000	GAGGTGGTGT
	1551	COGCGACGAC	CCCCTCCACT	TOGTOGTACA	30333003A0	03403333373	TGGGTCGGGC
20	24421	AGTOGACGGO	SATGCGGCGC	ACCCAGACGC	cacassioni	21,827,033,03	
	16641	0.340330000	2999030000	GCGACGGTCG	TGGTGGTGGC	1007775033	CCCGCGACCC
	5-5-5-	AGACGCCCTC	GATCCGGGCC	GGATOCGCGT	CGACG70330	000000000000000000000000000000000000000	GCGACCGAGC
	and the second	-00AT000000		AGTTOGGGGTA	GGAGCAGGAG		AGCGCGACGA
	26761		SOGTOCGGGG			AAUGUTGCGC	
~ -	16811	GGCGGGGCACC	STEETECAGG	GTGAGCGCTC	CGGCGACACA	3300303663	ATCTCSCCCT
25	16661	GGGAGTGTCC	GATGACGGCG	TOOGGGGGGTA		STECCACACG	GCGGCCAGCG
	26941	ACACCATGAC	GGCCCAGCAG			SCGGCGGGTC	ACCTCCGGGT
	27001	CGTCGAGCAT		TCCCAGCCCG	TGTGCGGGAT	CAGCGCGTCG	GCGCATTGGC
	27061	SCATOCTGGC	GGCGAACACC	GGGGAGGCCG	CCATCAGTTO	GACGCCCATG	CCGCGCCACT
	27121	GCGGTCCTTG	TOOGGGGAAG	ACGAAGACGG	TGCGCGGCTC	GODOCAGE	STOCCOGTGA
30	17181	CGACGTCGTC	STOGAGCAGO		GCGGGAACGT	COTACGCCTG	GCGAGCAGGC
	27241	CCGCGGCGAT	GGCGCGCGGG		GACGGGGGGG	GAGGTGCTCG	CGGAGTCGGC
	27301	GGACCTGGCC	GTCGAGGGCC	STGGCGGTCC	GCGCCGASAC	GGGGAGTGGT	STGAGCGGCG
	27361	TESCRATCAS	CGGCTCACCG	GGCTTCGAGG	CCGACGGCTC	crosscege	SSCTCCCCGG
	27421			ACCIGGGGGG	TGGTGCCGCT	GACGCCGAAG	GAGGACACAC
35		CCGGGTGGGC	TTOCAGCAGG	GTCTCGGGGC	AGGGCCGGGC	AT063T6A66	AGTTCGACGG
22	17481	0330303000	CGGGGGGGTCG		3037370330		
	07541	CECCECCET	CONSTOGACG	TGCGAGGACG		GTGCAGGGTG	CGCGGCAGGG
	27601	Taccaracca	1A.T.GGC6A.G6	ACCATOTYGA	TGAGAGGGGG	3808000000	GCGCCCTGAG
	17661		STIGGACTIC		GCAGCACCGG	3373733030	CCCTGCCGGT
		AGGT3300AG	CACCGCCTGT	GCCTCGATGG	GATOGOCCAG		GTGCCGTGCG
40	17781	CODOCACGO	GTCCACGTCC		GCCCGGCGTT	3300A36600	TGCCGGATCA
	27841	coddaracard	CGAGGGCCCG	TTCGGCGCCC	ACAACCCGTT	GGARGDACCG	TOCTGGTTGA
			CCGGACAACC	GCCAGCACAC	GGTGGCCGTT	GCGCTCGGCA	TOGGAGAGCO
	27961	TOTOGACGAT	CAGCACACCG	GACCCCTCGG	CGAAACCGGT	GCCGTCAGCC	GCATCCGCGA
	28021	LOGOTTOOL	GCGCGCGTCG	GGCGCGAGAC	cooggracia	GGAGAACTCG	ACGAAGCCGG
45	1002	TOGGET COM	CATCACCGTG	ACCCCCCCC	CCAGGGGGAA	CARACATTOS	COGGAGCGCA
7.7	20001	2003000000	GGCCTGGTGC	Acceptage.	GCGLCGLCGL	LOLOGOCGTO	TOGROCGTGE
	20141	3.3.5.10000	CTCCAGACCG	MICONOCAL CALL	TOTOGOGO	70101010103	200000000000000000000000000000000000000
	26201	CUBCUBBACC	CICCASACCG	DOGRACIAGO	MCMGCGGAGG	OBT COCCACO	0130101000
		13003370GC	GCCGAAACCG	CCCAGGTCGG	1GCCGMG.CC	GIACUGGICG	CACMARGUCGC
5 ()	28321	COATGAACAC	GCCGGTGTCG	CTTCCGCGCA			GCGLGIICUM
50	29391	GOGCOTOCCA	CGAGGTCTCC	AGGACCAGAC	GCTGCTGCGG	GICCATCGCC	AGOGCCTCAC
	25441	GOGGACTGAT	CCCGAAGAAC	GCCGCGTCGA	AGTCCGCCAC	CODESCEAGE	AAGCCACCAT
	28801	GACGCACGGT	CGACGTGCCC	GGATGATCCG	GATCGGGATC	GTACAGCCCG	TOCACGTCCC
	08861	AACCACGGTC	CGTCGGAAAC	GCCGTGATCC	CGTCACCACC	CGACTCCAGC	AGCCGCCACA
	16611	AGCCCTCCGG	CGACGCGACC	CCACCCGGCA	g00000TA300	CATOCCCACG	ATOGOCAACG
55	28681	3000000000	COSSACGGCC	GCGGTCGTGG	1306367033	CGATGCCGTC	CGGCCGGACA
	28741	2020000000	SAGCTTCGCC	GCGACGGCGC	neggeerden	GAAGTOGAAG	ACCGCGGTGG
	12001	2020220001	TACCCCCGTC	- 1000.0000000	#8866444600	TIRCOCCITTO	POPARTAGOS
	- : : : -		GAGTTCCTTG	HOLDEURGER		200200000000000000000000000000000000000	COSTOSTESS
	20001		CGCGGTGCAC	MACCOLOGICA		CACCEDANCC	SCGTCGTGGG
211	21921	- Denot AUGGC	DECESTEDAD	Tannaganan	mcomccccc	G1CCITITOS	000100000
60	at #61	LGGAGAGCCG	CGCGATCCGG	وافافاشطاعا والادا	1001000000	000000000000	2020000000

PCT/US99/22886 WO 00/20601

33

				_			
	1.2	0000000000	7903030A30	AGGGGGAGT	73273E3322	33000330703	30330GACCA
	2277	GCCCCCGGCTC	03A33A00000	21,000,000		CAGTOCOCCT	COGGOGGTOR
					2222222222		
	2317		0000703033			SACCSTCAGE	
	`		30000083372	A CONTRACTOR		1900000033	
5				772303337		2737003333	CTCCCGAGGA
-							- C.
	13041	0800000000	CGACGAGTAG	សំផល់សំពេកប៉ុន	1170262717	1373700763	eleneul
	19471	30A33T30CA	0000000000	A0000000000	3380387777	CTCGAGGGGG	T000000TGA
		30300073%3	GA000000100	TOGAGGACCC	100000101	CACGACGGCC	3734503367
	- 314 O L					JACJACCJ.,	
	13501	30000036577	3873033333	AGTACAGARA		2233703603	A007000A06
10	1921	contropeden	SACCTOSSES	COCCCACCA		300 307303 0	GROMGCATCA
• •			300.00.000	COTTOGACGA		**********	
	13641	30x3003303				amalani uti da	addhadaldd
		CGGRGCCRCC	0973A03A30	Aldersetat		~30 733A6 03	TDACCOSSCG
	1.30	2012200000	3600K3K033	000000077727		ACCCAGCACC	1,0000.00000
							BOBTOBACSA
			3373332737	GOGAGOAGOG	30.032332.		
15	11.4651	33A03AT003	90000000000	100000000000	132722200	TART003303	3003033003
		ADDOGAGACO	000000000000000000000000000000000000000	780A. J030A	93A00303T0	330 37 A0 C 33	TOGTOGGTGA
					00000AGTT0	3033376703	
	30001	GGAAGIGCTG	CACGGCGGTC	AGGACGCCGG			TOGAGOGGGG
	30061	CACCGCCGCC	GCCGTGCGCS	GGGAGGATCA	CONCETTORS	HACCGTCGGG	TOGTOGAGGC
	30121	GGCCGGTCGT	CGCGGTCGTG	GGCGGCAGCT	0000040000	SSCCAGGACC	SGGCGCAGCA
20							
20	30191	GGCCCGGAAC	GGCTCCCGTG	AllCGluAGoo	3330327333	napgedgeeg	ATBGTGGCGA
	30241	06660000600	GGTCTCGTCC	GCGAGGTGTA	000000000000000000000000000000000000000	PSTGACGGCG	ADDRIGHACOS
	30301	cogregoed	GGTGGCGTGG			COACCOARCC	T33TC2CCTT
						CAGGCCGTAC	0372033037
	30361	CCGCGGGGAG		GCGCCGAGCA	3030033373		
	30421	CGGCGAGCTG	TCCGTCGGCG	AGGGCCACTT	000000A3A0	SHIPPOTTOG	TOBGOCCAGA
25		0360606066	GCGGGGCAGC	GCGCGCCCGT	COGTOTACIO	000000000000000000000000000000000000000	AGACGGTCGG
 '	30461					2227000000	
	30541	CGATGTCGTC		GGCCGGGGCCG	1330333033	opino o onio	GGCATCTCCC
	30601	GCACGGCCGG	GGCCGTCCGC	GGGTCGGGGG	CGAGGATTCC	GTGCGCGTGC	TOGGTOCACT
		0000030030			00000000000	GCCCTCCGCC	ccssscscsc
	20001		310000000	19CM29313M		_	
		TCACCGTGAC			ACCOLOGUAG	03TGAGGGGG	STSTCCACGG
30	30781	TGAACGTGTC	GAGGGGGGGG	CAGCCGGCTT	0070000000	CCGGATCGCC	AGATOCAGGA
-			GGGCAGCACC		SCASSCASES	CGCCAGCGGA	TCGGCGGCGT
	30931	CGACCCGGCC	GGTGAGCACC	AGGTCGCIGG	TGCCGGGCAG	GGTGACCGCC	GOGGTCAGCG
	30081	CCGGGTGCGC	GACCGGCGTC	TGTCCGGCCG	3 3333333 3373	GCCCGCGGTC	TEGGTGCCGA
			GACCCGCTCG		103G0GGGT0	CGAGGGGGGT	3003603036
35	31061	GGTCGATGAC	CTTCGGCCAG	TOGACOGTGA	CGCCGTCGGT	GTGCAGCCGG	GCGAGCGCGG
	31141	TCAGGGGGGA	TOSCOGRATOS	TOGTCGGCGT	GCAGCATCGG	SATGCCGTCG	ADGAGTOGGG
			GTCCGGGGCCG	A M C M C C T C C T		amagmanana.	GCGACCTGTT
	31201	TUAGGOTOCG			Compositions	3133133333	
	31261	CCCCGAACCG	GACGGTGTCG	CGGACCTGTC		CTCCGGCGTG	GTGCAGGCGG
	3:50		CATCGGGATC		GGTROGTORG	SCTCTCCCCC	ACCTTGCGGA
10		100000000	on to occurre	*#00000000			0308330666
40	22302	ACTOCTOGAG	CHIDOSSICS	Alchoracia	na.Janness	3.300.3010	
	31441	TGAAGCGGCC	GAGCCGGGCC	GCGACGTCGA	GORCOGGCCTC	CINGTCACCG	GAGAGCACGA
	ខាន់ក្រ	TOGACGCGGG	COCCOMOLO	GCGGCGATTET	CCACGCCGTC	COGCAGCAGO	GGCAGCGCGT
	21201	CCCSTTCCGA	COCCITONICS	COCCCCTTCC	000000000	annangagaa	manamosago
	31201	CCCSTTCCCGA	CGCGATUAGG	60414400000	Ubbbbbbban	JOSSAGGGGG	.337.37.303
	31621	GGGCCCGTGC	GGACACCAGC	CTGCACGCGT	CCTCCAGGGA	CCAGACGCCG	GIBACGTACG
45	មាននគម	CGGCGGCCAG	openanealma.	GRATGGCCCA	0388366660000	CGGGGGTAGG	COCCACCCC
⊤ ~′	0.1001	CGAGCTGTGC	- creditation and a	3.00000R300	2011/01/01/02/02	22200000000	m: deddagaram
	31.41	CGAGCTGTGC	GCCG#G1262	AUCTOUNDED	benting a die	3330.00000	
	31361	CGTGGAGGTC	GAGCCCGGGG	GGCACGTCGA	magacardona	CACCTCGCGG	CSAGTGCGGG
	21621	CGAAGACGTC	optionoged	SOCISTICOST	ngoonsmann	REGEREGATET	GRECCCTGTC
	31001	JOHNOMUG. C	GIMOGDGGGG	3007310031	00000000000	pagagata	C T T T C C C T C C
	31921	CGGAGAAGAG	CCACACGAGG	CGGCGG'I'CCG	المودوديناف	and a specific transfer	6.01000100
50	21981	CGATCAGCGC	GGCCCGGTGC	GGGAAGGCCG	TGCGGGCGAG	CAGGGCCGCG	GCCACCGCGC
	22201	SCICGICCIC	2000000000	occapement.	000000000000	المتصديدة المتاسية	000000000000
	32641	SOLUSIUSIU		GCGWGQ LG 3G	53655775555	0.0000000	0000000000
	30101	CCTGCGGGGT	GCGTGCCGAG	AGCAGCAGSS	GiralGa.li.	rependancing	310000000
	30121	STTCSGGGGC	03350033333	TGGCTTTCGA	GGATGATGTG	ROCOTTOGTG	COGCTARCGC
		2110900000	010.0000000	00000000000		aggerages	0.000,000,000
		CGAAGGAGGA				3335580000	maanaamaan
55	32283	TGAGGAGTTC	GACGGCGCCG	GCCGTCCAST	CGACGTGTGA	JGACGGCGTG	: JUACUTSUA
	21241	GGGTGCGCGG	ALCOGRACIO	TECCECATES	CGAGGACCAT	CTTGATGACA	COGGCGACGC
	32342	000100000	omerce = =====	200000000000000000000000000000000000000	20000000000	2020120202	1000000000
	32401	2CGCGGCGGC	CTGAGTGTGG	CCGATGTTGG	MULLIUMGUUM	SUCCAGUAGE	7000000101
	37.223	CGTGATGCTG	CCCGTAGGTG	GCCAGTACCG	0076030377	GATGGGGTCG	CUCHAUL 40
	20001	TOCOCOTOCO	ATGCCCCCCCC	ACAGOGTOCA	CATCCGCCGG	GGTGAGCCCG	GUGTTGGUUA
(0	34541	GCGCCTGCCG	7.100300103	maamaaat aa	00000000000	CCCCCCCCACC	COSTRUCADA
60	32581	GUGUUTGUCG	am Fumululusu	100100000	0000011000	Caccatorate	2001,000.4.0

			Suggestion of the Control of the	GANCER TOEK	0080000000	SACRITOTES	COGRECOGOT
	1 - 		the property of the second	- 00-1-00-00-00-0		0000000000	CCGGTGCCAT
		in i		- Madhilandan - Amaga sagaa		- De Data Cammé. - Se pagadaga	TGGTGGGAGA
		and the state of the state of	Company of the contract of the		o ješti i pista autik Ogađana se od se	-amagoodug0 00000000000	
5			100000000000000000000000000000000000000	maadd in Lin	and the second second second	e umu umu umu e Ombobaca	3135535A30
*						22120212100	SACGMACACG
	32340	آريڪي ٿي ۾ شاه ٿو جي انهيماني عجم جي جي جي ج		- 33M, 00100M	- MMC US LA WER		00A0033A0A
	12 . 2	الرائي في الوالية والأنجاب المنا	7163313719	しゅいほうしだってるだ		i de a de a de a de a de a Compositor de a de a de	STOCCUTACC
		COTACAROTA		ARCACCCCGG	TOTOGOTT	and article of the	. c 0 3/3/3R/30R.
1	11111		TT0TA90907	7220A23A33		TAGACGCTGC	TGCGGGTCCA
ŀΟ	11111	7.30728333	277220006	organocida		202223220	GOCARDOOGG
	13241	120,430,430,	ACCAIGACGC	ADDGTTGAGG			GGATOGTACA
	23271	30000770A0	3770072073	0337003703	GAAANGGCCCC	DATOZZSTOR	T0A2000A0T
	22341	20%20%2000	COACAAGTOC	T003303800	OGROCOPAC:	1930A32033	CAGGCCATCC
	3341.1	2002000000	07/40000000	7777797233%	03300037731	030337A030	0000033763
15	3.3 1.4 1	T0000000000	3003000837	7007004367	33323323A3	0322732323	3738337367
	33341	CGARGACGAG	03TA305350	AGCGTCAGGC	0007030377	3992880008	TTGCGCAGTT
	23011	CORCCCCGT	CAGCGAGTCG	AAGCCCACTT	-000003AH000	1130303035	GCGATGGCGT
		3335970303	GTGGC0GAGC	Accepaacha	CGCTGGTAAG	HACGAGGTCG	AGCATGTCGC
	33721	0000030033	ABBTOCGGAC	G13050233h	033003302.0	383333333	AGGACCAGCA
20	35781	GGACCCCCTC	33A0838333	ACGGCGGCGA	GSTCGNGCCC	GATGGGGAAGG	AGCGCGGGCC
	33841	G377337373	CAGGGCCCCG	TCGAACA333	CGAGCCCCT:		ATCGGGGGTCA
	33301	T3203TT303	9003AT3033	GCCAGGTCGG	7000307083		Andodonoed
	33961	-00000T000A	CROTOCCOAG	GCGAGCGAGA	0330333332	0000030100	T300007030
	346.01	0000070000	GTOGAGGAAC	GCGTTGCCGG	TOGCGTRGTT	3300738300	GOGCCGCCGA
25	34081	ACOTEGOSGA	TATGGACGAG	TACAGGACGA		STOSHSHTOS	CGCGTCAGCT
	34141	COTOCAGGTO	CCAGGCGACG	TCCGCCTTGA	CCCGCAGCAC	GGCGTCCCAC	TGCTCCGGCC
	34201	GOATGGTGGT	CACGGCCGCG	TOGTOGACGA	TCCCGGGCCAT	GTGCACGACG	GCGCGCAGCC
	34261	30TG0303A0		ACTGCGGCCA		GTCGACGACG	TOGGOGGOCA
	34301	CGTACCSCAC	3033T03T00	TOCGGCGTGT	2GCCGG33C3	3006773036	GACACCACGA
30	34351	CGACCTCGGC		ACGGTGAGCA		SASSASSSG	CCGAGCCCGC
50				ACGGTCCCGC	CGCTCAGCGG		GTGGCCGCGG
	34441	CGGTGCCGCC	GGTGACGAGG CAGACGGGCC	GCACGCGCTG	TGCCGTCGGC	GACCCGGACG	TGCGGCTCGT
	34501	CGACACGCCG		GCTATGGCGG	CGGGCGTSAT	c1001000100	TOGATCAGGG
	34561	00000300300	GAGCCCGGCC	GTGTGCGCGCG		300000000000	GOTTOCTGCG
35	34681	0340303330		GCCACGATGA	TOCGGAUUAS GCCGGGATOS	26000A3390	GGCTCGGCGA
25	34651	CGGGATCGCC	GGTACGGGTG	AGCAGGTCGC	GGCCCAGCTC	0000010000	GCGCCGGGCG
	34741	SOCAGGTCTS	CACGGTGGTG		GGACCACGAC	033333730	TOGOCGTOGG
	34821	ASGTGCCCGG	GTCGCCGGGT	TOCACGGCCA	CGGGTGACGC	333CA023SC	ACCCAGGCGA
	34861	GCACGTCGGC	GAGGTACGTC	CASTCGGGGA	CGGCCGGCAE	3370A0330TG	TOGACGTOAA
175	34921		CONCINCATOR	TOGGGGTTCGG	CGATGCGGAC		DOGACGOGTT
40	74981	30A003613A	GCCGTGCTCG	TCCGTGGCGA		0810100333 0810100333	
	35041	CONGCAGONO	GCGCAGCGCG	GTCGCGGCGC	GCGCGTGGAT	GAGGGGGTGC	GACCAGGAGA
	35101	ACGCCAGCCG	GCGCCGCTCC	GGGTCCGTGA	AGACCGICCC	383333313C	70000000001
	35161	CGMSCAGCAS	GGGGTGCAGD	CCGTACCGGG	mecessars	013.103363	COCCOCCOCCO
	35001	ACGCSTAGGC	GCGGCCCTCC	CCCGTCCALA	TOGCOGIAN.	addecedance	magagagagaga
45	35281	ACGAGAGCGG	CAGCGCGTCG	TAGAAGCCGG	TCAGG redge		100000000
	35341	GCCAGTCCAC	gggCTCCGCC	GGACCGCUAG	16	DAGGGC1CCG	GICGUALIGA
	35401	GOGCCCAGGG	GCCCGTGCCG	GTACGGCTGT	GCAGACTCAL	LOMULGULGI	TOBGRESS I
	35461	CGGTTCCGAC	SETGGCCTGG	ATCTCCGTGT	CGCCGTUGUU	BRAGERERA	AUCUGUGGA
	35521	CGATGCTCAG	CTECGCGATC	TOOGGOGTGO	CGAGCCGGGC	TOUCUGUTTUG	GLGAGCAGII
50	35581	CCACGAGOGC	CGAGCCGGGG	ACGATGACCC	GGCCGTTCUAL	0.000000000	GUIDAGUUAGG
	35641	GCTGACGGCG	TACCGAGACA	CCGCGGGTGGC	CAGCGTGCCC	TestesTess	GCGAGGTCGA
	36,301	CCCACGAGCC	GAGCAGCGGG	TGGCCGGACG	TTCCCGCGGG	1100303703	ATCCAGTAGC
	35761	GGTCACGGCG	GARCSGSTAC	GTGGGCAGCG	GCACCACCCC	AUGUSTUGUG	AACGACCAGG
	35 321	TGACGAGCAC	GCCCCGGACC	CAGAGOGOGG	CGAGCGACCC	AGTGAAGCGG	TODAGGCCGC
55	35.851	0070300103	CCGCAGTGTG	CCGGTGACGA	CCGTATGCGC	ATGCICGICG	AGCGTGTCCT
	35941	CCAGTGCGGT	GGTGAGCACG	GGATGCGCGC	TGACCTCGAC	GAACGCGCGG	TATCCGCGGT
	36001	CCGCCAGGTG	GCCGGTCGCG	GCGGCGAACC	GAACGGTGCG	GCGCAGGTTG	TOGTACCAGT
	36061	AGG0G300#0	10006660066	TOCAGOCACG	CCTCGTCCAC	GGTGGAGAAG	AACGGGACGT
	36121	COGGOGGGGG	COGRAGIGATE	CCGGCGAGAG	CGTCGAGCAG	CGCGCCGCGG	ATCGTTTCGA
60	36181	CATGCGCGGT	GTGCGACGCG	TAGTCGACGG	CGATCCGGCG	GCCGCGGGGG	GTGGCGGCCA
-							

				30030A0030	Janian dan d	00A0000 0 37	
	36301	73303A3373	CAGGCGCCCG	40000A0A033	-03903103 <i>N</i> A	3703603630	GGCACCGAGA
	34381	2247600322	0130006660	AGTTTTTTTTT	13A03A3T03	GOTGOGOACO	SCSACGACCT
		770000000		ASTACTORS	113A033X321	The state of the state of the state of	~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~
-	36421		GTCCAGGGTG				
5	3/481	A3T33303A0	GACCGCGGCC	G3G3CGACCC	0073783A1:	CURCAGOTOC	
	36541	002202200	GRACGACGCG	GGCTGCACGA	CATOGACOCA	3000MA0000	-330G0TCCGG
	36801	30000110000	SATSACGTCC	AGCAGGTTTT	ATCRESTER	70000000mgc	GCCGTGGCGC
	36461	ACTOGGGGAG		AACACGGGGGT	222222222		SSSATGCCSS
						- Profession	
		000A00003A	11121132223	3 33A A0303A	ACACGACACA		ACSTOGGOGG
10	3.00 5.1		39000000000	ACTICGGCAC	CA0333333AA	000000000000000000000000000000000000000	70703333003
	3,243.5	2080080000	0000100000	ATGGCCGTCC	GGGTGGTGGT	3A323A3733	CCGACCGCGG
				300%307373			
	112		ma.aMacaaa		oppounted.		
	36861	GGGCCGACAT	0333308430	A0070077333	3080033071	J. 15999.	GOGGACACGG
		3730393030	3303333333	coegoctock	GORDGRORD	3732773373	CCCCTGATGC
15	37125	CURACUARDOA	CACACCCCA	caccassia	307736738.	333345333	TOACTGCGGT
10				TCCCACTCGA	1313006388		ADGTGCAGOG
	3 29-	30A30A3003	SATSTOSCOS				
		T0000000A3	GACGCCGTGC	CGCATCGCCA	. 35005		SESADGCCGS
	37261	0000030000	GGTGTGGCCG	ATGTTTTAATT	TGAGCGAGG	TAT CASCASC	GGATGCACGC
	37321	3770003001	GTAGGCCACT	TGCAGGGCCT	GGGCCTCGAC	3333703000	AGACGGGTGC
20					CACCOGGCGC		TOGGCGAGCG
20	37351	1337300373	TGCCTCCACG	GOGTCGACGT		02493009309	
	37441	CACCCTGGAT	GACGCGCTGC	TGCGCAGGCC	CGTTCGGGGAA	PRACESCOOS	TTCSACGCGC
	37501	COTOGGASTT	GACCGCGGAG	CCGCGCGCACCA	GCGCCAGCAC	ABBGTGGCG	TBBCGGGTGG
	37561	COTOCOACAG		ACCAGGAGAG	0390900013	GBCGAAGCTC	STECCGTCCS
	37621	0337376066		GCACGGCCGT	C333GGCSA3		CGGGAGAACT
25	37681	CGACGAACCC	GGTCGTCGTC	GCCATCACCG	TGACACCCC	UNCONGGGGG	
	37741	CCCCCGAGCG	CAGCGACCGC	GCGGCCTGGT	GCAGCGCCAC	CAGOGACGAC	GAACACGCCG
	37801			CCCTCCAGAC	CGRAGTRGTA	08A8A60080	CCGGAGAGAA
		TGTTGACGGT			CGCCCAGGTC		COGTAGCCCT
	37861	COCTOGTOGG	CSTGCCGGTC				
	37921	GGGTGRACGO	GCCCATGAAT	ACGCCGGTGT	CGCTGCCGCG	GACGCTTTCG	GGUAGGATGU
30	37981	COGCTCGTTC	SAACGCCTCC	CACGACGCTT	CGAGGACCAG	ADGCTGCTGC	GGGTCCATCG
2.0	38041	CCAGCGCCTC		ATCCCGAAGA	ACGCGGCGTC	SAAGTOGGOG	GCGCCGGTGA
					CGACCGCGTC		TOGTAGAGOG
	38101	GGAAGCCGCC	GTGACGCACG	GAAACCTTGC			
	38161	CGGCGAGGTC	CCAGCCGCGG	TOGGOGGGGA	ACTOGGTGAT		CCGGAGTCGA
	38221	COAGCOGCCA	CAGGTCCTCC	GGTGACCGCA	COCCACCGGG	CATCCGGCAC	GCCATGGCCA
35			CGGCTCGTTC	CCCGCCACCG	TOSSTGCSGS		GCCGGAGCGG
22	38281	CEATOSCOAG		0203007020		9000A0000G	
	38341	07/36/36/06/3	CTCACCCCGG	CGTTCCTCAT	CCA36CGGGG		
	38401	GGTGGTCGAR	GACGGCCGTC	GCGGAGAGCC	GTACCCCCCC	23777723333	AGGCTGTTGC
	38461	33FA0003H0	ACCOCTGAGO	GAGTOGATGO	CRAGGTCCTT	G##10300ST0	GTGGGGGTGA
				CCGAGGAGGG	CGGCCGTGGC	DECACACACS	ATGGCCAGCA
	38521	TOTOGGASSO	STOGGCGTGG				
40	33581	OGTCACGATC	GCGGTCGCGG	TCGCGGTTGC	GGTTGTCCTC		GOGATGOGGO
	38641	GCTCGGTCCG	CTGCCGGACG	GGCTCGGTGG	GAATCGCCGC	GACCATGAAC	GGCACGTCCG
	32701	CGGCGAGGCT	COCCTCGATG	AAGTGGGTGC	CCTCGGCCTC	GETGAGOGGO	CGGAACCCGT
	20701	00000000000	encorecement	CCCTCCTCT		GAGGGTGCTG	
	30/67	TraceCNPCC a	0100000100	GCG1CG1CAM	0110 00001	07.00010010	0700007000
	38821	ACATGCCCCA	GGCGATGGAG	GTGGCGGGTTT	GGCCGAGGG.	GTGGCGGTGG	U JUNGOU . D
45	38881	CGTCGAGGAA	GGCGTTGGCG	GCGGCGTAGT	TTCCTTGTCC	sassersees	AGGACGGCGG
. •	300/1	occacomos:	GTAGAGGACG	AAGEGGGGGA	GGGGTTGGTT	TTGGGTGAGG	TGGTGCAGGT
	35391	2335661368	000000000000000000000000000000000000000	AAAEAATAAT	Component	GOGGTOGGGG	anglaggggg
	39001	GCCAGGCGGC	GTTGGGTTTG	Mesmeetteee	CG3.30.383	3103110000	0107000001
	39061	CGAGGATGCC	STOGTOGAGG	GTGGCGGCGG	TGTGGAAGAU	SSCGGTGAGG	55.150000m
	39101	Tattgggggala	GGTGGTGGCG	AGTTGGTGGG	GGTCGCCGAC	GTTGGCAGGGG	AGGTGGGTGC
50	30101	2222222	2722222222	consenses	1020616372	GGTGTGGGGG	TGGTTCAGGT
50	20201	133000103.	3103300001	************	*********	SECTIONS	3 n 4 m 4 m m C C i 3
	22547	530GGGGGAG	SATULUI BUU	WGGGTGCGG	macacacaca.	GATGATGATG	######################################
	39301	GGTTGAGGGG	GGTGGTGGTG	GGTGGGGTGG	TUGTUTGGAG	GGGGGTGAGG	
	34351	SSAGGGTGTG	GTGGGTGAGG	CGGAGGTGGG	GGTGGTCGAG	GETGGCGAGT	TGGGCCAGGG
	2240.	0010000100	amediacomed	woodenmout.	makaccaaat	-2033TGGGGG	TOTTCGTTCT
		- CANDOUGAU -	0.00000100	10001111111	2200000000	croscroscro	onstreamer.
55	39481	333066 T 363	GUTUAGGCCG	فأفأء فافالموالافا		0.000100.0	9.01334038
	39541	TGAGGGTGTG	GTCGGTGGTG	GTGAGGTGGT	GTTGCAGGGC	SGTCAGGACG	
	3925	117000000000	3733376337	ATGTCCTCCG.	GGTCGTCGCG	ATGGGCGGCG	GTGATCAGGA
	3 2 3 2 2		2000.000.	200000000000000000000000000000000000000	0000000000	3ACCGCGAGC	CACTOCARCO
	39001		0000m0010m		0000010000	Green anno	CISTOCACCC
	39721	SGAGCGGGTT	CGGCCCCGAC	GGGGTGTTGG		CAGCACCAGO	mannesses es
60	39781	ACACGACAGG	ACGCCCATCC	GGGTCGGCCA	CGCGCACGGC	GACGCCGGCC	

		:5895333A:	3000A00303	33300035553	TOBOOTTORE	0030A03000	STOCAGGAGA
	100	Accorac	331777703003	222222222	3312000000	3732 A3 33300	GOGTOGAGOA
	3 4 10 1			0000000000	00000000000	2722726332	AGGGGGAGGT
	*				and the second s		SCOSMOCCCT
5	41 61		33087000330	A0110313A1	AGRACOCCGA	GAC3703AC3	3006056603
	11.111	T330033033	000000000000000000000000000000000000000	AAGGGTTGRO	0008400000	33A33TAT30	333 373733 3
	40301	3337783637	G0030TG30G	1000000100	A307500000	00001106011	TGCGCGTGGA
	40061	0337070033	00000007200	GCCTCATCGG	200000000000000000000000000000000000000	POTORCOGRO	ACATOCACOG
	40301	0130331020	000000000000000000000000000000000000000	AG000338ATT	7371380370	77 7770707	ACCCCGCAAC
10	41341		A009300033	ATGACCAGGT	one are record		AGDAGAACCG
	40441	13000000000	COCOTGATCA	GOOMBOOMSS	SATSOGTAGE	CARTORGATO	CGGCCGGTGA
	10501	37262746375		T0000000000	STOOTSTOR!	7073300AG0	ATGGGATGCG
	41841		DA30003300	3033707331	733 73337 A34	3920300000	ABCCAGTACC
	41821		GAMODOGTAG	GTGGGGAGAT	CORSONSES	COACCCCC	GGTTCARCCA
15	47651	200730000A	3700822220	GCACCCAGAC	- Tanagaar .	7300AACGCC	CCCAGCCAGC
٠.	10.741		ACCSTCACCA	GTCCGCMAC.	:GOCACOST	0000000000	TCCATCGCCG
	40800	30730773076	000AT333CA	OTOCACTOCA	03240400337	1000700/30	TOCGCCACCG
	40561		0000ACAGGG		7,00078008	71.80000 0 05	TOCACCGGOT
	40921		0000000000	ACGGTCGACC			CCGCCGGAAA
20	40921	BERRESEA	TACCTCAGCC	AGTTCGTCGT	COATOGOCTO	38337 385 60	GTGTGGGAGG
-17	41041	COTACTOMS	00000000000	CGCACCCGCA	0000770730	CTCATACCC	GCCACCACCT
	4 - 0 4 -	0010020000	0606ATA060 0606A0366700	0000000000	00005.10.10.100 000T0GAAG0	DOGROCATTA	CGCGCCGCGA
	41161	TOCACACACC	CTCGACCAGA	0001000000	00000030AA	COCCACCONA	
		-100m0m0m00 -00000000000	CASCCSCSCS	GOGATOACCC	GACTGCGCAA	2000AC0A00	CSGCGGCGT
25	41221	COTOCAGGOT		GCCACACACA	000000000		GAGTGTCCGA
رد س	41251		GAGGGCTCCG CGGCACGACC	COATGCGCCT	gggAcAgogc	SSCONGGCTC	
	41341	COACAGOGTO		ACCTCCACCC	GCTCCGCCAC	ATCCGACCGC	GACAACATCT
	41401	CCCAGCTGGC			ADGCCCGCGC	ACACTCCTCC	ATACGAGCCG
	41461	CCCGCACATO	CCASCCCSTG	TGCGGCAACA	CGCCCATGCC	CACCCACIGG	GCACCCTGCC
20	41521	CGAACACCGC	GGAACGGTCC	ATGASTTCCA	CCACCGCCAC		CGGGCATCAC
30	41581	CGGGGAAGAC	GAACACCGTA	CGCGGCTGAT	CACGCTCACG	CACCAACCCC	TGCGCGACCG
	41641	COAGCAGCAC	CSCACGGTGA	CCGAAGACAG	ACCCCTCCAG	CCCCTCCACC	TGCCCCCGCA
	41701	CGGCCACATC	CACCCACCC	COGOGOAGAT	ACCCCACCAA	ADDADGACCA	
	41761	GACTCACCTC	ACCACGAGGG	GACACCGGCA	~~~~~~~~~	GTTCGTACCG	CTCACCCCGA
3.5	41331	CACGCGACGG	COCAGGAACA	CCCTCCAGGA	-castomessa		GCCTCGGTGA
35	41881	ACGACGACAC	ACCCCCATGO	COTSCOOGAT	CATGCGACGA		ACGTGCAGCG
	41941	GCAGCTCCAC	DGCACCGGCC	GACCACTOCA		GATGAGAGGG	GCGACACCCG
	40001	TOTTCGGCGC	GATCCCATGC	0005100005	TGACCATCTT		GGCGTGTCGC
	42061	0830000000	CGCATGACCG	ATGTTTTTA	TGACCGAACC	GAGGTAGAGC	ASCCGCGTGC
	30121	SGTCCTSCCC	GTAGGCCGGG	ようのいしつりょうと	GOGGGTGGAT	0.000.00000	
40	42161	0337300373	CGCCTCCACC	ACGTOCACAT	0330330303	CASTCOGGOG GGACAGTCCG	TTGACCAACG TTGGAGGCAC
	12241	CCTGCCGGAT	0A06030700	TGGGCGACGC	2011 - 000000		
	42301	COTOCTOCTT	CACCGCCGAG	CCGCGGACGA	CCGCGAGAAC	GGTGTGCCCG	TTGCGCTCGG
	42361	COTOGORGAG	CCGCTCCAGC	ACGAGAACGU	CORRECT CT C	magagaaamaa	GICCCGICCG
	42421	CCGCGTCGGC	GAACGCCTTS	CACCGTGCG1	DUSGUGAGAS	1000000000	TOGGRAGAMCI
45	42481	COACGAGOTO	TGCGGTGTTC	GCCATGACGG	TGACACCGCC	GROCES CONC.	MAGGAGCAC:
	42541	cadassucas	CAGTGCCTGT	GCCGCCTGGT	GUAGGGCGAC	CAGUGACGAC	GAGCACGCCS
	42601	TETEGACEST	GACCGCCGGG	CCCTGAAGTC	CGTACACGTA	CUAGAGGCG	CCGGACAGGA
	42661	CGCTCGTCTG	CGTCGCCGTG	ACACCGAGCC	CGCCCAGGTC	CUGGUUGACG	CCGTAGCCCI
	42721	GGTTGAACGC	SCCCATGAAC	ACGCCGGTGT	CGCTCTCCCG	emboulfGTCC	SCHOGATGC
50	42781	CGSCGTTCTC	GARCGCCTCC	CAGGAGGTCT	CCAGGATCAG	6300100163	GGGTCCATCG
	42841	CCAGCGCCTC	STICGGACIG	ATGCCGAAGA	ACGCGGCGTC	GAAUUUUGGG	CUGGULAGGA
	42901	ATCCGCCGTG	GOGTGTCGTG	GAGCGGCCGG	CCSCGTCCGG	GTCCSGGTCG	TACAGCGCGT
	12961	ICARCOTOCAL	30000055703	GTGGGGGAAGT	CGGTGATCGC	CTCGGTACCG	GCGGCGACGA
	43021	G00000ACA3	STECTOOGSO	GAGGGGACCC	03006660A3	TOGGGAGGGG	ATGCCGAUGA
55	43031	magagaaaaa	GMCGCCGGAG	CCGAGGGTCT	GGGGGTCGC	GGGTGCCGCT	GTCGCGGAGU
	43141	000000000000	-nacadagaan	GCECGCGGEG	TSGGGTGGTC	GAACGCGGTT	GANGUUUUU
	43101	0000000000	~~~~~	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	TOSTGAACTC	-GACGOTGGTG	AGOGAGTIGA
	. 2061	2,5,5,5,0,0,0,0,0,0,0	0000110070	-dggTcccc55	AGCAGTGTCC	3336000630	mound jour
	43333	CARROBORES	GOTGTOGGGG	ACCAGGTOGA	GIAGTACGIC	CTCCCGGCCC	CCACGGGCCG
60	43381	CGGCGAGGCG	GTTCGCCCAC	TOOTETTOOG	TGGCGTCGGG	CTCGGCCGGT	CCGGTCAGTG

	11141	Section of the contract of the contract of	35.55000000				00007#0000
	43501	DOGG BASSAT				00000000000	2000000000
			SATSTACOAS				0000000000
	43661			3304305050		0073000ACT	
ς.	13461			703303000		PRO MOGROS	
	13661		0000007.700	100001000	7337080371		0101031300
	4.2	GGAGCAGGCC	330323 331 6	raggegraat	000000000000000000000000000000000000000	- inseatioses	3-3-5-3005
	43801	CGRTCGGAGG	COCCACGOTG	AGGACCATCT	TGGGGGTGTG	11.13303133	CTCATCCACG
	43661	CGRACGCGTC	CCGCGCACGG	CGGATGTCCC	ACGGCTGCAC	0000 0000	CACAGCTCAC
	45921	00000TCGAA	CAGGTCGAGG	AGCAGTTOGA	GGATOTOSSS	07.03090909	GGATTIACST
10	43981	CGGCCAGGTC	GAMOGGOTGO	TGGGCGGCGT	330 33 AT3T0	3370773000	ATOTOGACGA
	44541	A200000000	CGGTGCGAGC	AGGCCGATGG	ACGCCTCCAC	SAGTTOACTG	GTGAGCGAGT
	44111	TORGOROGAD	GTOGRECOGG	653/4106737	00000AM0000	000000000000000000000000000000000000000	GAGTTCGCCA
	44181	CATGGTCGGT	STOGRAGOOG	TCGGGGTGTA	SCAGGTGTTS	TTTTGGGGGA	0733033133
	44021	COTACACCEC	GGGGGGGAGG	TGGCGGGGGA	11030511331	100077773000	ACACCGCCCG
15	44351	7000000000	GACCAGGACC	Trongonia:		30000000000	ACGAGGGGGG
	44241	K0CA38033T	GOGGAACACG	ATGGGCAUSS	1978.	GREERACGAC	CATCCCCCT
	44451	GBATCCGTGC		ceamodecea	00%000000000000000000000000000000000000	31.3%ACGCG	COCTOCACUA
	44461	GACCGAMICAC	GCGGTCGCCG		datogadgat	COSTCOGACT	TOGGTOROGA
	44521	1300030303		ATCTCGCCCT		101000GAG0	GOGATOAGOA
20	44561	CGTCGCGGAA		GCGGCGCGGA	CBTCGATGCG	3A000000000	GCGGCCAGGG
	44641	9030999999	ACGTCGAGCG		GAGGT03033		AGG0GGG0GG
	447.1	3090490300		GTCGGCAGGG		030000TA00	AGCOGGGGGA
	44761		G00G3000G0		GGGGTTCGCC	BAGCGAGGCC	GCGGCGGGGA
		- CGAGGTCGTC				TOCOGGTTCG	G0GG007G30
25		GGCGCAGCGC	CTCGTCCCAG				G003GG003A
- -J	44881	CGCCCACCGC					GGCAGGTIGC
	44941	GCCGCTCCCA	C 1 C 3 C 3 C 3 C 3 C 3 C 3 C 3 C 3 C 3	ATANCACARCA	CCTCGCCACT	300337 00	
	45001	COGGCAGCC	GACCAGILLOG	CCCCCCCTCCS		1303313036	
	40001	TGACGTGCCA	CEUCAGOCACC	COCCOCTOON		00000000000	
30	90161	TOMOGRAGOOMA	GALCICGICG	CCCCCCCCCC	COCLAROSIC	STAGAGCACG	
30	40161	GGATCGCCTC	GGCGGGGACG	CCCCCCCCCC		GACACCGGCC	
	45241	CGAGGACGGG	GIGCGGGCGG	CCCGCCGCGG	TOGGGGTCCCG	GOSGATGATO	
		CGACGGTCTC	GATUTULUG	#CMCCCCMCM	200000000		COGGTGCGGT
		cccggccggT				GGCGTGGTAG	
2.5		ACCAGCCGTC	CACGAGUACC	TOBOCCOCOCOCOCOCOCOCOCOCOCOCOCOCOCOCOCOC	corposedes	TBOCACGTCG	
35		GGCTCGGCCC	GCTCGCCCAC		30A03G3CAG	COCCCACCAC	
		CCGGGGTCGAC	GAACCGCAGC	GALAGGGGG			
	45601	GOGCATCOTO	CAGGGTGTTG			CTCGGTGCAG	
		CGAGCAGGGG	CACGCCGRAC				GGCGAGGTGG
4.2.	45721	ATCCGGCGAC	CAGTGCCACG	0004600000	GAGCCCCCGG		ACGGCGCCGA
40	15761	GGAGGTAGCG	GTACATOGTO	GGCACGCCGA	COMPONED A	GCTGGAGTGT	TCGGCCAGGG
	45.541	CGTCGAGGAC	GTCACGCGCG	ACGAAGCCGC	CCAGGATACG		CCGACCGTGA
	45901	GGACGGCGAG	CAGGCAGAGG	TGGTGGCCGA	GGCTGTGGGAA	TAGOGGGGGG	
	45961	GTTCGTCGTC	CTCGGTCAGC	CGCCAGGACG	GCACGTCGCA	GTGCATCGCG	GACCACAGGC
	46921	CGCTGCGCTG	TGCGGAAAACC	ACGCCCTTGG	GACGGCCGGT	GGTGCCGGAG	GTGTAGAGUA
45	46081	TCCAGGCGGG	TTCGTCCAGG	CCGAGGTCGT	CGCGGGGCGG	GCACGGCGGC	TOGGTOOOGG
	46141	CGAGGTCCTC	GTAGGAGACG	CAGTCCGGTG	CCCGGCGCCC	GACGAGCACG	ACGGTGGTGT
	46201	CGGTGCCGGT	GCGGCGCACC	TGGTCGAGGT	GGGTTTCGTC	GGTGACCAGC	ACGGTCGCGC
	46261	CGGAGTCCGT	CAGGAAGTGG	GCGAGTTCGG	CGTCGGCGGC	STOOSGGTTG	AGCGGGACGG
	46321	CGACGCCGGC	GGCGCGGGGCG	GCGGCGAGGT	AGACCTCGAT	BETCTCGATC	CGGTTGCCGA
50	46301	GCAGCATICGO	GACCOSGTOG	CCGCGGTCGA	CGCCGGACGC	GGCGAGGTGT	CCGGCGAGCC
	46441	gaddagadaaa	GAGCCGGAGT	TGCGTGTACG	TORCGGCGCG	TTGGGAATCC	GTGTAGGCGA
	46501	TOOGGTOGOO	GCGTCGCTCG	GCATGGATGC	GGAGGAATTC	GTGCAACGGC	CGGATTGGTT
	36551	conciona	CATGGLAACA	CCTTTCTCTC	GACCHACCGC	ACARCAGCAC	GGAACCGGCC
	48601	ACGRGTAGEC	GCCGGCGACG	CTAGCAGCGT	TTTCCGGACC	GCCACCCCT	GAAGATCCCC
55	46681	CONTROLL	degenoone	GGACGCTCAT	CTAGGGGGTT	SCROSCATAC	CGCCGTGCGT
	36731	AATTGCCTTC	creataine	ATGCCGGACG	COAGGGAAGG	STESAGGCGT	TGTCCATATC
	16563	marcacacac	0000000000	GCTTCGAGAA	GACCGGATTA	COGGACCTCC	AGGGTGACGA
	22225		GGCCCCETCG	AGCACGGCAC	03 G CCACACO	GACGTGTCGC	TGGTGGACGG
	16971	TARTORCOCAS	indegreemen.	ACACCACGAC	-CCGTGAUGAC	GAGGCGTTCA	CCGAGGLU.5
60	12001	GCACGCACAG	0600000000	AGTCCGGCAT	GGACAACGGC	ATCGCCTGGG	CCCGCACCGA
00	ಇಲ್ಲಿ	uunvaahund	55555.00				

				7 3 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5		ABRITACGOCO	1000010000
	1 1 2		in a significant constraint of the significant constraint co			COSTATOCCO	*********
	i vis		modele e	in a seed to do a seed for a seed to be a seed		38.3903 07 33	
	100	The Control of the Co	Tradicial instance	and the ring of the country of the	ancarmos		
_	-	1.775 - 1.1 - 1.1				1338703800	. Jajoneon.
2	4	3708830337			30359357.5	30011303191	
	47.242		STOCESATCA	mumi bumumn			accac abled
	47401	on concount		CCCCGGTTTT		ACCTGGCTGG	
	47461		CONSTRUCTION	CCGCGACGGC		GGACACCGAA	
	37121		ROCCCCCC	GCGAGGTCGC		ATGGCCCGGG	TOATOGGOGG
10	47551	39AG/ACCT3	0000000A00	GCGTCCAGCG			TGGACCACCT
	47841	CARGGTOTAC	3700800300	GCGAGGATCT		0900366777	GCGCCGCACG
	40001	cerements	A000000000	TOGCOCTITI		ATABOCCECS	AGGATCTGCT
	47761	CONCERMANC	GAMGGCATGG	TGGCGTGACA	ATACCCGGTA	A46300050	GACGCTGCGC
	47501	0.7000000000	10000844346	AAAGAAGAGO	GTCACCGCAC	AGCGCGGCAG	CCCGGTCCTT
. 2	47881		TADAGCGGCG	GATOTESTE	OTO ASC 17	BBACCCGGA	GAGCAACGCC
	47941	77 77 77 77 77 77 77 77 77 77 77 77 77		ACGCCTGCGC		Addooccaga	CCTGGAGCGT
	48001			GCGCCACGAG	GOGTTGCGGA	COGTOTTOGA	CACCGCCGAC
	48061	9900A00000	700A000000	G0TT000300		nicrecocca	capacagaaa
	48121		AGGACGCCGC			Toboogcoc	STTCGACCTC
20	45161	300M00GB30		ggggggggggg	ATGGGGGTGG	GTGACGACGA	CCACGITCIC
		300 300333 003		CGCCGGCGAC	GGCTGGTCGT	***********	CCAACATGAA
	45241	crosciscos		200000000000000000000000000000000000000	Equipment of	0000004400	
	48301	000000000000000000000000000000000000000	WOLUDOUS COM	CGCCGCC73G		AACTCACCGG	CGCCGGACTG
	48361		TGGCCTACTG	GCGCGAGCAA		0000000000	
25	48431	GACAGGGGTC		GCCGGTCGCC		CSGSCATGGC	
25	48481	JUDDUUMUU	GTCCCCGCCC	CGCGGTCGTC		GOGACTOOGG	
	48541	0000000000		CTTCCAAGCG		GGCAGGCGGG	
	45601	TTCATGACCC	TGCTGGCGGC	GGCGAACCGT		CGTACGAGGG	
	48661	Breergeres	GCACGCCCGT			GCGATCCGTC	
3.0	48721	ATGTTCGTCA	-	GCTGCGCGGC	CACCOCTTCC	CCCACGCCGA	
30	43781		GCTGCCGGGC	CACGACCALG	GACGCG11CG	COCHCIA CCC	CCIGOCOLIC
	48841		TOGAACTOGT			CGGTCAACCC	
	48901			GCGCGACGCG		CGCTGCCCGG	
	48961	GRACOSTICO	GCACCGGACG	CTGGTTCACC		TOGAATTOCA	
	49021		GCGCGCTGAC	CGGCGAACTS		STGCGCTGTT	
3.5	49081		GGTTGCTGGA			AGGCGGTCAC	
	49141	GACGTACGGC	TGTCGCGGCT	GCCGGCCGGC		CGGCAGCGCC	
	49201	TOGRACGACA	CGGCGCGGGA	COTGCCCGTC	GACACGCTGC		GGCCCGGTAC
	49261	GCCGCACGCA	00000000000	CGTGGCCGTC		.,	CACCTACGCG
	49321	CASCIGGACO	33033303AA	COGCCTCGCG		900000000	CACCGCCACC
40	49381	SGCGACCTGC	TOGGGATOTG		GGCGCCGACC	TGATCGTCGG	CATCGTGGGG
	49441	ATCCTCLAGG	0999090030	TTATGTGCCG	CTGGACCCCG	ARCHICCTCC	GGAGCGCACG
	49501	CETECTTECE	TGGCCGACGC	GCAGCTGACC	ACGGTGGTGG	CGCACGAGGT	
	49561	CGGTTCCCCC	ATGTGCCGCA	CGTGGTGGCG	TTGGACGACC	CGGAGCTGGA	CCGGCAGCCG
	19671	concioned	030033333	CGAGCTGGAC	CGGGACAGCC	TUGUCTACGC	GATCTACACG
45	19681	TOOGGGGGGG	coggoaggeo	GAAGGCCGTG	CTCATGCCGG	GTGTCAGCGC	CGTCAACCTG
	49741	CTGCTCTGGG	AGGAGCGCAC	GATGGGT130	GARCCGGCCA	GCCGCACCGT	CCAGTTCGTG
	19901	LOGOCOACGE	magachachd	GGTGCAGGAG	ATCTTTTCCG	CGCTGCTGGG	CGGCACGCTC
	49961	GEGERACCESC	CGGACGAGGA	GCGGTTCGAC	CCGCCGGGAC	TOGCOCGGTG	GATGGACGAA
	19921	Cannoamma	22222222	CGCGCCGACG	GCCGTACTGC	GCGCGCTGAT	CGAGCACGTC
50	19981	GATCCGCACA	GOGROCE	cgccgcccts	CGGCACCTGT	GCCAGGGCGG	CGAGGCGCTG
5 ()	50011	Importable of	addaggmmada	CGAGCTGTGC	CGGCACCGGC	CCCACCTGCG	CGTGCACAAT
	50101	CACTACGGTC	coccociiis	CCAGCTCATC	ACCGGGTACA	CGCTGCCGGC	CGACCCCGAC
	50131	GOGTGGCCCG	-00000000000000000000000000000000000000	GATOGGOOOG	CCGATCGACA	ACACCCCCAT	CCATCTGCTC
	20101	GACGAGGCGA	magagagaga	#CCCCTCC3#	amgnegggge	AGOTOTGOGT	COCCGGCGTC
55	000223 00001	GGCCTCGCCC		GGCCCGTCCC	annance.	COGREGEE	GGTGCCGGGA
ا د ر	2000 2000 2000	GATGCGGTCS	013331801	C1464440000	20000000000	TGGCCCGCCG	CGCGCCCGAC
	55391 55455	G30GA3CT33	000000000000000	CHIGINGEIC	CAPCAGAGA	2000000000	CATCCGCGTC
	50401	9906A00133 9AA0036313	mATTUULUSS	COUNTRICARY	020202020		000000000000000000000000000000000000000
	50461	- 3FA0036316 - T006160603	mamilwanana	0018311339		0101070007	#CCGGTGGCC
10	50521	T000T00003 -3500330A23	ACONTOCO CO	COCCOCCCC	CTRCCCCCCCC	0207330030	CCGGCTGCCC
60	5056l	3500330A23	ocoheani".	1900 a 600 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	010000000	0,72,20000	3000010000

	1.00		adadaaquiji	1272255555		CASSAGGAS	nthataenan
	era de la compania d La compania de la co		and the second second			000000000A0	2222222222
		AAGGTGGATT	33.3.3.3.3.	3033A033T3			232002001
	5. 5.						op.ednoord composition
ž			an can Salahan menan	325		2 2 2 2 2 2 2 2 2 2 2 2 2	mana amaman
Š			ili di distribili di				130000000000
		GACGGUCGGA		3/21/2/2/2/2/2/2/2/2/2/2/2/2/2/2/2/2/2/2		muduunduu,	
	1 1 1 1 1	707000AT03	10000110000	9989FR 5030	223722777		ACAGGAACAG
	41001		CACACGGCTC	3073010222			CCCGTACGGG
		TT00000T31	BOSSSCARCT	COMOCOCOGAN	9090113443	000000000000	CCGGATCGCC
10	51131	GOGGGGGAGG	AGDOGOTOCG	GA000003TT0	_030GAT 0303	FACAGGTCGT	0033303000
	51041	GOTOCOGTOC	BBBBBBBBBT	GGTTCCGGTG	C CCCCTCCCCCC	ACGTCGACGC	cacaarccaa
	51301	GTTGGCCCACC	GGGAGCTGAC	COGGCCGTTC	GROOT 3373A	ACGGGTCGTT	GCTGCGTGCC
	flact	STOCTOCTOC	CEDEDEDECE	CGAGGATCAC	3730730730	TGATGCTGCA	COMPORTOGOC
	51421	GGTGACGGAT	GGTCCTTCGA	ccreerages	CGGGAGTTUT	COGGGGACGCA	ACCGGACCTT
15	51481	TOSTOTOCCO	ACACGGACGT	GGCCCGGTGG	02400334373	000000TGAT	COCGGCCAGG
* = "	51541	GAGAAGGACC	GGGCCTACTG	G0GC033033	0133333333	COECCOCC	GGAGCTGCCC
	51661			ACCGACCGGG	223222777	TOTOGACGCT	CANGGACACC
		G000T00G30	00000000000	GGTCGCGGAC	20001100100	COMCGTTGCA	
	51661	GCCGTCCTGG	CGGCACGCCG			2021001025	G0T0GT0G0G
30	01 11		TOSCOCOTOST	CGTGGCGGAG	Adda ve are a		
20		ACCCCCTTCC	CGGACCGGGG	GTACGCCGGG		TOATOGGOTT	
	51841	GTCCTCGCGC	TGCGCCTCGA	coressesse	7.020.20.00	TOCCOCASOT	
	51901	3TGCACACCS	CGATGGTGGG	COCCACACAC	DAGDA (300)	TGCCCTACTC	
	51961	GCGAGGACC		3006930000			COCCOTOAGC
	52021	GCGGAACTGC	SGCTGCCCGG	シスカム ひしんしんしゃ	3#30033777		CGAGACCGTC
25	52081	GACGAGATGA	COGGCGAACT	GTCGATCAAC	2727723223	ACGGTCGCAC	
	50141	scastaatec	ACGATGCCGC	Gargarage	7373007.003	TUBACGATTT	GCTCACCCGG
		GTGGAGGCGA		caccacasac	GACCTCACCG	TACGCGTCAC	
	50061	SEAAGOGAGE	AGCCATGCCC	GAGCAGGACA	AGACAGTOGA	GTACCTTCGC	TOGGOGACCG
		CGGAACTCCA			CCGCGCACAG	CGAGCCGTTG	
30	50301	000000000000000000000000000000000000000	CCGGCTGCCC	GGCGGGGTCG	CGTCGCCGGA		
50	J = J = J = 1	GONTAGCCIO	CC00C100CC	0000000	00.00		
	E24/1	t cmccccmcc	CONCCCCATC	racecemmac.	CCACGGGGGGG	GGGCTGGGAG	ACCACCGCCG
	52441	AGTCCGGTGG	CGACGGCATC	ACCGCGTTCC	CCACGGACCS	GGGGGGGGAG	ACCACCGCCG
	52501	ACGGTCGCGG	CGGCTTCCTC	ACCGGGGGGGG	CCGGCTTCGA	CGCGGCGTTC	TTCGGCATCA
	52501 52561	ACGGTCGCGG GCCCGCGCGA	CGGCTTCCTC	ACCGGGGGGGG ATGGACCCGC	AGCAGGGGGT	CGCGGCGTTC GGCCCTGGAG	TTCGGCATCA ACCTCGTGGG
3. -	52501 52561 52621	ACGGTCGCGG GCCCGCGCGA AGGCGTTCGA	CGGCTTCCTC GGCGCTGGCG GCACGCGGC	ACCGGGGCGG ATGGACCCGC ATCGATCCGC	AGCAGCTTCGA AGCAGCGCCT AGACGCTGCC	CGCGGCGTTC SGCCCTGGAG GGGCAGTGAC	TTOGGOATOA ACCTEGTGGG ACGGGGGTGT
35	52501 52561 52621 52631	ACGGTCGCGG GCCCGCGCGA AGGCGTTCGA TCCTCGGCGC	CGGCTTCCTC GGCGCTGGCG GCACGCGGGC STTCTTCCAG	ACCGGGGCGG ATGGACCGGC ATCGATCCGC GGGTACGGCA	AGCAGGGGGG AGAGGGTGGG TOGGGGGGGA	CGCGGGGTTC CGCCCTGGAG CGCCAGTGAC CTTCGACGGT	TTCGGCATCA ACCTCGTGGG ACGGGGGTGT TACGGCACCA
35	52501 52501 52621 52681 52741	ACGGTCGCGG GCCCGCGCGA AGGCGTTCGA TOCTCGGCGC CGAGCATTCA	CGGCTTCCTC GGCGCTGGCG GCACGCGGGC STTCTTCCAG CACGAGCGTG	ACCGGGGCGG ATGGACCCGC ATCGATCCGC GGGTACGGCA CTCTCCGGCC	COGGOTTOGA AGCAGOGOOT AGACGOTGOO TOGGOGOGOA GOOTOGOSTA	CGCGGCGTTC GGCCCTGGAG GGGCAGTGAC CTTCGACGGT CTTCTACGGT	TTOGGOATCA ACCTCGTGGG ACCGGGGGTGT TACCGGACACCA CTGGAGGGTC
35	52501 52561 52621 52631	ACGGTCGCGG GCCCGCGCGA AGGCGTTCGA TCCTCGGCGC	CGGCTTCCTC GGGGCTGGCG GCACGCGGGC STTCTTCCAG CACGAGCGTG GGTCGACACG	ACCGGGGGGG ATGGACCGG ATCGATCGGC GGGTACGGCA CTCTCCGGC GCGTGTTCGT	CCGSCTTOGA AGCAGOGOT AGACGOTGO TCGGGGGGGA GCCTCGGGTA CGTCGCTGGT	CGCGGCGTTC GGCCCTGGAG GGGCAGTGAC CTTCGACGGT CTTCTACGGC GGCGCTGCAC	TTGGGCATCA ACCTCGTGGS ACGGGGGTGT TACGGCACCA CTGGAGGGTC CAGGCCGGGC
35	52501 52561 52621 52631 52741 52801	ACGGTCGCGG GCCCGCGCGA AGGCGTTCGA TOCTCGGCGC CGAGCATTCA	CGGCTTCCTC GGGGCTGGCG GCACGCGGGC STTCTTCCAG CACGAGGGTG GGTCGACACG CTCCGGCGAA	ACCGGGGGGG ATGGACCGC ATCGATCGGC GGGTACGGCA CTCTCCGGGC GCGTGTTCGT TGCTCGGTCG	CCGGCTTCGA AGCAGCGCT AGACGCTGCG TCGGCGGCGA GCCTCGCGGT CGTCGCTGGT TCCTGGTGGT	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC CTTCTACGGT CTTCTACGGT GGGGGTGGAC CGGGGGTCACG	TTOGGCATCA ACCTCGTGGG ACGGGGGTGT TACGGCACCA CTGGAGGGTC CAGGCCGGGC GTGATGGCCT
35	52501 52561 52621 52631 52741 52801	ACGGTCGCGA GCCCGCGCA AGGCGTTCGA TCCTCGGCGC CGAGCATTCA CGGCGGTCAT	CGGCTTCCTC GGGGCTGGCG GCACGCGGGC STTCTTCCAG CACGAGCGTG GGTCGACACG	ACCGGGGGGG ATGGACCGGC ATCGATCGGCA CTCTCCGGGCA CTCTCCGGGCC GCGTGTTCGT TGCTCGGGAGC	COGGOTTOGA AGCAGOGOOT AGACGOTGGO TOGGOGGOTA COTOGGOTGGT COCTGGTGGT AGGGGGGOTT	CGCGGCGTTC CGCCCTGGAG GGGCAGTGAC CTTCTACGGT CTTCTACGGT CCGCCTGCAC CCGCCTGCAC CGGCGTCACA	TTOGGCATCA ACCTCGTGGG ACGGGGGTGT TACGGCACCA CTGGAGGGTC CAGGCCGGGC GTGATGGCCT CCGCCCTCA
35	52501 52501 52621 52681 52681 52801 52801 52801	ACGGTCGCGG GCCCGCGCA AGGCGTTCGA TCCTCGGCGC CGAGCATTCA CGGCGGTCAC AGTCGCTCCC CGCCGCGGCGCAC AGGCGTTCGC	CGGCTTCCTC GGGGCTGGCG GCACGCGGGC STTCTTCCAG CACGAGGGTG GGTCSACACG CTCCGGGGAA GTTCGGGGGAC GGAACCGGCT	ACCGGGGCGG ATGGATCCGC ATCGATCCGC GGGTACGGCA CTCTCCGGGCC GCGTGTTCGT TCCTCCGAGC GACGGCACCG	COGGOTTOGA AGCAGOGOTT AGACGOTTOGOTA COCTOGOTTACT COCTOGOTTACT AGGGGGGGOTT GTTTOGOCGA	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC DTTCGACGGT DTTCTACGGT DCTCGACGGT CGGCGTCACG CGGCGTCACG GGGGTCACG	TTOGGCATCA ACCTCGTGGG ACGGGGGTGT TACGGCACCA CTGGAGGGGGC CAGGCCGGGC GTGATGGCCT GCGCGCTGCA GTCCTGATCG
	52501 52501 52601 52601 52601 52601 52601 52601 52901	ACGGTCGCGG GCCCGCGCGA AGGCGTTCGA TCCTCGGCGC CGAGCATTCA CGGGGGTCAT AGTCGCTCGG AGGCGTTCGC AGGCGTCGC	CGGCTTCCTC GGCGCTGGCG GCACGCGGGC GTTCTTCCAG CACGAGCGTG GGTCGACACG TTCCGGCGAA GTTCGCGGGAA GTTCGCGGGAC GGAAGCGGCC	ACCGGGGGGG ATGGACCCCC ATCGATCCCC GGGTACGGCA CTCTCCCGGCC GCGTGTTCCT TGCTCCCGACC GACGGCACCC GAGGGCACCC	COGGOTTOGA AGCAGOGOOT AGACGOTGGO TOGGOGGOTA COTTGGTGGT AGGGGGGTTT GTTTTGGCGA GCCACCGGTT	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC DTTCTACGGT DSTCTCTACGGT DSCCTGCAC CGGCGTCACG GGGGTCACG GGGGTCACG GGGGTCACG GGGGTCACG GGGGTCACG GGGGTCACG	TTOGGOATOA ACCTCGTGGS ACCGGGGTGT TACCGCACCA CTGGASCGGCC CACCCGGGC GTGATGCCT GCGCCTGCA GTCCTGATCC GTCCTGATCC GTCCTGATCC
	52501 52561 52681 52681 52681 52681 52681 52881 52881 53881 53881	ACGGTCGCGA GCCCGCGCA AGGCGTTCGA TCCTCGGCGC CGAGCATTCA CGGCGGTCA AGTCGCTCCC CGGGGAGAGCT TCGAGAAGCT	CGGCTTCCTC GGCGCTGGCG GCACGCGGGC GTTCTTCCAG CACGAGGGTG GGTCGACACG GTTCGCGGAA GTTCGCGGAC GGAAGGGCT CTCCGACGCC CCAGGACGGCT	ACCGGGGCGG ATGGATCCGC ATCGATCCGC GGGTGTCGT TGCTCGGGCC TTCTCCGGAGC GACGGCAACG GACGGCAACG GACGCCAACG	COGGOTTOGA AGCAGOGOTTAGA TOGGOGOGA GOOTOGOGAT COTTGGTGGT AGGGGGGTT GTTTTGGGGGA GOOACCGGG GCCACCGGG GCCACCGGG GCCACCGGG GCCACCGGG GCCACCGGG GCCACCGGG	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC DTTCTACGGT DTTCTACGGT DCTCCTACGGT CCGCGCTCACG CGCCCCGAC GGGGTCCGGC GCCGCGGTC GCCGAACGGGTC GCCGAACGGGGTC GCCGAAACGGG	TTOGGCATCA ACCTCGTGGG ACCGGGGGTGT TACCGCACGACA CTGGAGGGTC CACCGCGGGCGGGGGGGGGG
	52501 52561 52681 52681 52681 52681 52881 52881 53981 53181	ACGGTCGCGG GCCCGCGCA AGGCGTTCGA TCCTCGGCGC CGAGCATTCA CGCCGTCAC AGGCGTCAC TCGAGAAGCT CCGCCGTCAA AGGCGTCAA AGGCGTGAT	CGGCTTCCTC GGCGCTGGCG GCACGCGGGC STTCTTCCAG CACGAGGGGG GGTCGGACAGG STTCCGGGGAA STTCCGGGGAC GGAAGGGCT CTCCGACGGC CCAGGACGGCT ACGGCAGGCC	ACCGGGGCGG ATGGACCCCC ATCGATCCCC GGGTACCGCC GCGTGTTCGT TGCTCCCCCA TTCTCCCGACC GACCGCCAACC GACCGCCAACC GCCTCCAACC CTGGCCAACC	COGGOTTOGA AGCAGOGOTTAGAGAGGGGGGGGGGGGGGGGGGGGGGGGG	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC DTTCTACGGT DTTCTACGGT GGGGTCACG GGGGTCACG GGGGTCCGGC GCGGGGCGGAC GCCGGAACGGG CCCGGAACGGGCCGAACGGG CCCGGAACGGGAC CCCGGAACGGG	TTOGGCATCA ACCTCGTGGS ACGGGGGTGT TACGGCACGA CTGGAGGGTC CAGGCGGGGC GTGATGGCA GTGATGGCA GTCCTGATCG GTCCGGGGTT CCGCCGGGGTT CCGCCGCAGG GTGGACGCG
	52501 52501 52601	ACGGTCGCGG GCCGGCGCA AGGCGTTCGA TCCTCGGCGC CGAGCATTCA CGGCGGTCAC AGTCGCGTCGC	CGGCTTCCTC GGCGCTGGCG GCACGCGGGC GTTCTTCCAG CACGAGCGTG GGTCGACACG TTCCGGCGAA GTTCGCGGGAA GTTCGCGACGCC CCAGGACGCC CCAGGACGCC	ACCGGGGGGG ATGGACCCCC ATCGATCCGC GGGTACGGCA CTCTCCGGGCC GCGTGTTCGT TGCTCGGCGACC GACGGCACCG GACGCCAACG GCGTCCAACG CTGGCCAACG	AGCAGCTTOGA AGCAGCGCCT AGACGCTGCC TCGGGGGCCGA GCCTCGCCGA CGTCGCCGA AGGGGGGCCT GTTTCGCCGA GCCACCGCGC GCCACCGCCA GCCACCGCCA GCGACCCCAT	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC DTTCTACGGT DTTCTACGGT DSCCTGCAC CSGCGTCACG GSCGTCACG GSCGCGAC GCGGACCACACACG CGAGGCACAC	TTGGGCATCA ACCTCGTGGG ACCGGGGGTGT TACCGCACCA CTGGAGGGGC CAGCCGGGC GTGATGGCCT GCGCGTGCA GTCCTGATGG GTCCGGAGG GTCGCGAGG GTGGAGGCG GTGGAGGCG GCGCTGCTGG
40	52501 52561 52681 52681 52681 52681 52881 53881 53981 53161 53161 53221	ACGGTCGCGG GCCCGCGCA AGGCGTTCGA TCCTCGGCGC CGAGCATTCA CGGCGGTCAC AGTCGCTCCC TCGAGAAAGCT TCGAGAAAGCT TCGAGGACAA AGCGGTGAT TCGAGGGCCA	CGGCTTCCTC GGCGCTGGCG GCACGCGGGC GTTCTTCCAG CACGAGCGTG GGTCGACACG GTTCGGGGAA GTTCGGGGGAC GGAAAGCGGCT CTCGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC	ACCGGGGCGG ATGGACCCSC ATCGATCCSC GCGTGTTCGT TCCTCCCGGC TTCTCCCGASC GACGGCAACC GACGCCAACC GCCTCCAACC CTGGCCAACC ACCASGCTSC GACACCCCTG	COGGOTTOGA AGACGOTGCG TOGGOGGCGA GCCTOGGGGA CCTTOGGGGA AGGCGGGCT GTTTOGGCGA GCCACCGGGA CCGGACCGCC CCGGACCGCA TGCTGCCGA TGCTGCCGA TGCTGCCGA	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC DTTCTACGGT DSTCTGCAC GSGCGTGAC GSGCGCGAC GSGCTCCGGC GCCGGAC GCCGGAC CCGCGGAC CCGCGGAC CCGCGGAC CCGCGGAC CCGCGGAC CCGCGGAC CCGCGCGAAC CCGCGCGAAC CCGCGCGAAC CCGCGCGAAC CCGCGCGAAC CCGCGCGAAC	TTOGGCATCA ACCTCGTGGG ACGGGGGTGT TACGGCAGGGTC CAGGCCGGGC GTGATGGCA GTGCATGGATGCA GTCCTGATGCA GTCCTGATGCA GTCCTGATGCA GTCCTGATGCA GTCCTGATGCA GTCCTGATGCA GTCCTGATGCA GTCCTGATGCA GTCCTGATGCA GTCCAACATCG
	52501 52561 52681 52681 52681 52681 52881 53881 53981 53161 53161 53221	ACGGTCGCGG GCCCGCGCA AGGCGTTCGA TCCTCGGCGC CGAGCATTCA CGGCGGTCAC AGTCGCTCCC TCGAGAAAGCT TCGAGAAAGCT TCGAGGACAA AGCGGTGAT TCGAGGGCCA	CGGCTTCCTC GGCGCTGGCG GCACGCGGGC GTTCTTCCAG CACGAGCGTG GGTCGACACG GTTCGGGGAA GTTCGGGGGAC GGAAAGCGGCT CTCGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC	ACCGGGGCGG ATGGACCCSC ATCGATCCSC GCGTGTTCGT TCCTCCSGC TTCTCCGGASC GACGGCAACG GACGCCAACG CTGGCCAACG ACCAGGCTAACG ACCAGGCTAACG ACCAGGCTGG	COGGOTTOGA AGACGOTGCG TOGGOGGCGA GCCTOGGGGA CCTTOGGGGA AGGCGGGCT GTTTOGGCGA GCCACCGGGA CCGGACCGCC CCGGACCGCA TGCTGCCGA TGCTGCCGA TGCTGCCGA	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC DTTCTACGGT DSTCTGCAC GSGCGTGAC GSGCGCGAC GSGCTCCGGC GCCGGAC GCCGGAC CCGCGGAC CCGCGGAC CCGCGGAC CCGCGGAC CCGCGGAC CCGCGGAC CCGCGCGAAC CCGCGCGAAC CCGCGCGAAC CCGCGCGAAC CCGCGCGAAC CCGCGCGAAC	TTOGGCATCA ACCTCGTGGG ACGGGGGTGT TACGGCAGGGTC CAGGCCGGGC GTGATGGCA GTGCATGGATGCA GTCCTGATGCA GTCCTGATGCA GTCCTGATGCA GTCCTGATGCA GTCCTGATGCA GTCCTGATGCA GTCCTGATGCA GTCCTGATGCA GTCCTGATGCA GTCCAACATCG
40	52501 52501 52681 52681 52681 52681 52981 53981 53101 53101 53221 53281	ACGGTCGCGG GCCCGCGCGA AGGCGTTCGA TCCTCGGCGGTCAT AGTCGCTTCGC AGGCGTTCGC AGGCGTTCGC TCGAGAAGCT TCGAGGAGAAGCT TCGAGGGTCAA AGCGGGTGAT TCGAGGCCCA CCACCTACGG	CGGCTTCCTC GGCGCTGGGG GCACGCGGGC STTCTTCCAG CACGAGGGGG SGTCGACAGG CTCCGGGGAA GTTCGGGGGAC GGAAGGGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC	ACCGGGGGGG ATGGACCGC ATCGACCGC GGGTACGGCA CTCTCCGGGC GCGTGTCGT TCCTCCGAGC GACGGCAACG GACGCCAACG GCCTCCAACG ACCAGGCTAG GACACCCCTG GACACCCCTG GACACCCCTG GACACCCCTG	COGGOTTOGA AGACGOTGOG TOGGOGGOGA COTOGGOGGA COTOGGOGGA COTOGGOGGA COCACCGOGG COGGACCGOG TGOTGCOGGA TGOTGCOGGG GTGTCATCAA	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC CTTCTACGGT DTTCTACGGT DGCGCTGACG GGGGTCCGGC GGGGTCCGGGC GCGGACGGGAC CCGGGGGAC CCGGGGGAC CCGGGGAC CCGGGGGAC CCGGGGGAC CCGGGGGAC CCGGGGGAC CCGGGGGAC CCGGGGGAC CGAGGCACAG CTCGCTGAAG GATGGTCCTC	TTOGGCATCA ACCTCGTGGG ACGGGGGTGT TACGGCACGACA CTGGAGGGTC CAGGCCGGGC GTGATGGGAGGTC GTGATGGGAGGTT GCGCGGGGTGATGG GTCGAGGGGGGGGGG
40	52501 52501 52681 52681 52681 52681 52681 538101 53261 53261 53261 53261 53261 53261	ACGGTCGCGG GCCCGCGCGA AGGCGTTCGA TCCTCGGCGCC CGAGCATTCA CGGCGGTCAC AGTCGCTCCC TCGAGAAGCT TCGAGGAGAT AGCGGGTGAT TCGAGGCCCA CCACCTACGG GCCACACCCA	CGGCTTCCTC GGCGCTGGGG GCACGCGGGC GTTCTTCCAG GACGAGGGGG GGTCGAGGAGGGAA GTTCGCGAGGGC CCAGGAGGGC CCAGGACGGC	ACCGGGGGGG ATGGACCCCC ATCGACCCCCCCCCCCCC	COGGOTTOGA AGACGOTGOG TOGGOGGOGA GOOTOGOGTA COTTOGGOGA AGAGCGGOGA GOTACGGOGA GOCACGGOGA GOCACGGOGA GOCACGGOGA GOCACGGOGA GOCACGGOGA TGOTGCGGGG TGOTGCGGGG GTGTCATCAA	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC CTTCTACGGT CTTCTACGGT CGGCGTCACG GGGGGTCACG GGGGGTCGCGGGCGAC CCGGGGGGGAC CCGGGGGGAC CCGGGGGGAC CCGGGGGAC CCGGGGGAC CCGGGGGAC CTCGCGACGCCCCC CTCGCACGTC	TTOGGCATCA ACCTCGTGGG ACGGGGGTGT TACGGCACCA CTGGAGGGTC CAGGCGGGC GTGATGGGAGGTC GTGATGGGAGGTC GTGATGGGAGGTT CCGTGGAGGTT CCGTGGAGGTT CCGTGGAGGTT CCGTGGAGGGGGGGGGG
40	52501 52501 52681 52681 52681 52681 526981 53981 53101 53261 53261 53261 53261 53261 53261	ACGGTCGCGG GCCCGCGGA AGGCGTTCGA TCCTCGGCGCCCCCCCCCC	CGGCTTCCTC GGCGCTGGGG GCACGCGGGC GTTCTTCCAG GACGAGGGGA GTTCGCGGGAA GTTCGCGAGGAC GCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC GCAGGACGGC GCAGGACGGC GCAGGACGGC GCAGGACGGC GCAGGACGGC GCAGGACGGC GCAGGACGGC GCAGGACGGC GCAGGACG	ACCGGGGGGGGA ATGGACCCCC ATCGGCGCCCCCCCCCC	COGGOTTOGA AGCAGOGOTT AGACGOTTGO TOGGOGGOTT COTTGGOGGA AGGGGGGOTT GTTTOGGOGA GCCACCGGGT GGGGACCGGGT TGGTGGTGGG GTGTCATCAA LLLLCGCGGGC GGCCCTGGCC	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC CTTCTACGGT CTTCTACGGT CGGCGTCACG GGGCGCGAC GGGCGCGAC GCGGCGGAC CCGGCGGAC CCGGCGGAC CCGGCGGAC CCGGCGGAC CTCCCTGAAG GATGGTCCTC CTCGCACGTC CGAAACCGAC	TTOGGCATCA ACCTCGTGGG ACGGGGGTGT TACGGCACCA CTGGAGGGTC CAGGCGGGC GTGATGGCA GCGATGGGCC GACTGGACGC CGCCCACGC
40	52501 52501 52602 52602 52602 52602 52602 53101 53261	ACGGTCGCGG GCCCGCGCGA AGGCGTTCGA TCCTCGGCGCCCC CGAGCATTCA CGGCGGTCAC AGTCGCTCCC TCGAGGAAGCT TCGAGGACACCA AGCGGGCGCA CCACCACCCC CCGGCGCCCCC CCGGCGCCCCCCCC	CGGCTTCCTC GGCGCTGGGG GCACGCGGGC GTTCTTCCAG GACGAGGGGA GTTCGCGGGAA GTTCGCGGGAC GCAGGACGGC CCAGGACGGC CCAGGACGGC GCAGGACGGC GCAGGACGGC GCAGGACGGC GCAGGACGGC GCAGGACGCC GCAGGACGCC GCAGGACGCC GCCAGGACCCCC GCAGGACCCCC	ACCGGGGGGGGAACCGACCGACCGACCGACCGACCGACC	COGGOTTOGA AGCAGOGOTT AGACGOTGOG TOGGOGGOTT COTTOGGOGA GOTTOGGOGA GOTTOGGOGA GOCAGOGGOT GOTTOGGOGA GOCAGOGGOT GOTGTOGGOGA TGOTGGOGGA TGOTGGOGGO GTGTCATCAA LLLAGGOGGT GGCCTGGGO GCACCAACGG	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC CTTCTACGGT CTTCTACGGT CCGCGGTCCACG GCCCCACGGCGAC CCGCGCGAC CCGCGCGAC CCGCGCGAC CCGCGCGAC CCACATCATC	TTOGGCATCA ACCTCGTGGG ACGGGGGTGT TACGGCACCA CTGGAGGGTC CAGGCGGGC GTGATGGCA GCCATGGGC GCCATGGACGC CGCCCACGGC CTGGAAAAGCC
40	52501 52501 52621 52621 52621 52621 52621 52621 53101 53261 533401 53521 53521	ACGGTCGCGG GCCCGCGGA AGGCGTTCGA TCCTCGGGGGTCAT AGTCGCTGCG CGCGGGGGTCAT AGCGGGGGGTCAA AGCGGGGGGTGAT TCGAGGAGAGCT TCGAGGAGAGCT TCGAGGAGAGCT TCGAGGAGCGA AGCGGGGGGGGGG	CGGCTTCCTC GGCGCTGGGG GCACGCGGGG CACGAGGGGGAA CTCCGGGGGAA CTCCGGGGGAC CCAGGACGGC CCAGGACGCC CCAGGACCGCC CCAGGACCGCC CCAGGACCCCC CCAGCACCCCCC CCAGCCCCCCCC	ACCGGGGGGGGAACGGGGGGGAACGGGGGGGGGGGGGG	COGGOTTOGA AGCAGOGGOT AGACGOTGGO TOGGOGGGOT COTTOGGOGA AGGGGGGOTT GTTTOGGOGA GCCACGGGG GCCACGGGG GTGTCATCA GTGTGCTGGOG GTGTCATCA GCCCTGGCC GCACCACGGC GCACCAACGC GCACCAACGC CACCGGACA	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC CTTCTACGGT DTTCTACGGT DGCGCGCGAC GGGGCGCGAC GGGGCGAC CCACACCCCC	TTOGGCATCA ACCTCGTGGG ACGGGGGTGT TACGGCACCA CTGGAGGGTC CAGGCGGGC GTGATGGCAT SCCCGTGCA GTCCTGATGG GTCCTGATGG GTCCTGATGG GTCCTGATGG GTCCAACATCG GCCATGCGGC GACTGGACGG CGCCCACGGC CTCGAAAGCC CCGCTGCTGC
40	52501 52561 52621 52621 52621 52621 52621 52621 53261 53261 53561 53561	ACGGTCGCGG GCCCGCGGA AGGCGTTCGA CGAGGATTCA AGGCGGTCGGCGGTCAC AGGCGGTCAC AGGCGGTCAC AGGCGGTCAC AGGCGGTCAC AGGCGGTCAC AGGCGGCGCGT CCGGCGGCGCGT AGGCGCGCGT AGGCGCGCGT AGGCGCGCGT AGGCGCGCGT AGCCCGGCGCGCGT AGCCCGGCGCGCGT AGCCCGGCGCGCC	CGGCTTCCTC GGCGCTGGGG GCACGCGGGG CACGAGGGGG CTCCGGGGAA CTCCGGGGAC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGCC CCAGGACGCC CCAGGACGCC CCAGGACGCC CCAGGACCCCC CCAACTCCTC CTCCCTCCCTCC CCCCCCGAA	ACCGGGGGGGGAACGGGGGGGAACGGGGGGGGGGGGGG	COGGOTTOGA AGCAGOGGOT AGACGOTGGO TOGGOGGOTGGOT COTTOGGOGGA AGGGGGGGOTT GTTTOGGOGGA GCCACCGGGC GCGCGGGCG GTGTCATCAA LLLLCGCGGTC GCCCCTGGCC GCACCAACGC GCACCAACGC GCACCAACGC CACCGGACAC CACCGGACAC	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC CTTCTACGGT DTCTACGGT DGCGCTGAC CGCGCGCGAC GGGGCGAC CCACGCGACAG CTCGCACGCCCCC CCACACACCCCC CCACACACCCCC CCACACACCCCC CCACACCCCC	TTOGGCATCA ACCTCGTGGG ACGGGGGTGT TACGGCACCA CTGGAGGGTC CAGGCGGGC GTGATGGCTG GTGATGGCT GTGATGGCT GTGATGGCT GTGATGGCT GTGATGGCT GTGATGGCT GTGATGGCGC GTGATGGGCG GTGATGGGCG GTGATGGGCG GTGATGGGC GTGAACATCG GCGATGGGC GACTGGACGG CTGGAACGC CTGGAACGC CTGGAACGC GCGTTCCTCG GCGTTCCTCG
40	52501 52501 52621 52621 52621 52621 5262 5262 5262	ACCCCGACCA TOTOGGOCG	CGGCTTCCTC GGCGCTGGGG GCACGCGGGC STTCTTCCAG CACGAGGGGG SCTCGAGGAGGGT CTCCGAGGGGC CCAGGAGGGC CCAGGAGGGC CCAGGAGGGC CCAGGAGGGC CCAGGAGGGC CCAGGAGGGC CCAGGAGGCC CCAGGAGGCC CCAGGACGCC CGAACTCCTC CTCCTCCTTC GGCCCCGAA CACCCCGAA	ACCGGGGGGG ATGGACCCCC ATCGACCCCCC GGGTACGCCA CTCTCCGGCC GCGTCCCCCC TTCTCCGACCC GACGGCAACCC GACGCCAACCC GACGCCAACCC CTGCCACCCCCC GCCGTCGCCC GCCGTCGCCCCC GCCGCCCCCC GCCGCCCCCCC GCCGCCCCCC	COGGOTTOGA AGCAGOGGOT AGACGOTGGO TOGGOGGGOT GOTTOGGOGA GOTTOGGOGA GOTTOGGOGA GOCAGOGGOT GOTTOGGOGA GOCAGOGGOT GOTTOGGOGA TGOTGGTGGO GOCAGOGGO GOCAGOGGO GOCAGGOGGO GOCAGGGOGA GOCAGGGAGA CACAGGGAGA CACAGGGAGA TOGGGGAGAC TOGGGGAGAC	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC CTTCTACGGT DTTCTACGGT DGCGCTGCAC CGGCGTCACG GGGGTCACG CGCGCGGGGC CGAGGCACAG CTCCCCCCCCCC	TTOGGCATCA ACCTCGTGGG ACGGGGGTGT TACGGCACCA CTGGAGGGTC CAGGCGGGC GTGATGGCA GTGATGGCA GTGCTGATGG GTGCTGATGG GTGCTGATGG GTGCTGACGGC GCGTGCACGGC GCGCTGCACGC GCGCTGCACGC GCGCTGCTGC GCGCACCACGC CCGCTGCTGC GCGCACCACGC CCGCTGCTGC GCGCACCACGC CCGCACCCACGC CCGCACCCACGC CCGCACCCACGC CCGCACCCACGC
40	52501 52501 52621 52621 52621 52621 52621 52621 53621 53261 53261 53261 533401 53581 53581 53581 53581 53701	ACGGTCGCGG GCCCGCGGA AGGCGTTCGA TCCTCGGGGGTCAT AGTCGGTCGC TCGGGGGTCAT AGTCGGTCGC TCGGGGGTCAT AGCGGGGCGT CCGGGGGCGT ACCCGGGGGCGT ACCCGGGCGCGT ACCCGGGCGCGT ACCCGGGCGCGT ACCCGGGCGCGT ACCCGGGCGCGT ACCCGGGCCGT ACCCGGGCCGT ACCCGGCCGCGT ACCCGGGCCGT ACCCGGGCCGCGT ACCCGGGCCGCGT ACCCGGGCCGCGT ACCCGGGCCGCGT ACCCGGGCCGCGT ACCCGGCCCG ACGACAACCC	CGGCTTCCTC GGCGCTGGGG GCACGCGGGG GCACGCGGGAA GTTCGCGGGGAA GTTCGCGGGGC GCAGGACGGC CCAGGACGGC GCAGGACGGC GCAGGACGGC GCAGGACGGC GCAGGACGCC GCAGCCGCAC CCGAACTCCTC GCCCCGCAA CACCCCGCAA	ACCGGGGGGGGAACGGGGGGAACGGGGGAACGGGGGGGG	COGGCTTOGA AGCAGOGGCT AGACGCTGCGA GCCTGGCTGGT CCCTGGCTGGT AGGCGGGGGG GCCTGGCGGG GCACCGGGGG GCACCGACG GCACCGACG GCACCGACG GCACCGACG GCACCGACG GCACCGACG GCACCGACG CACCGACG CACCGACGC CACCGACAC CACCGGACAC CACCGGACAC CACCGGACAC	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC CTTCTACGGT DTTCTACGGT DGCGCTGCAC CGGCGCGCGC CGCGCGCGCGCC CGCGCGCGCCC CGCGCCCCC CGCGCCCCC CGCCCCCC	TTOGGCATCA ACCTCGTGGS ACCGGGGTGT TACCGCACCA CTGGACGGGC GACCCGGGC GTGATGCCT GCCCGCTGCA GTCCTGATCG GTCCTGATCG GTCCTGATCG GTCCTGATCG GTCCTGATCG GTCCTGATCG GTCCTGATCG GTCCTGATCG GTCCTGATCG GTCCTGCTGC GCCTGCTGC GCCTGCTGC GCCTTCCTCG GCCTTCCTCG GCCTCCTCG GCCTCCTCG GCCTCCTCG GCCTCCTCG GCCTCCTCG GCCACCCACT AACGCCGCC CCCACCCACT AACGCCGCC
40 45 50	52501 52501 52621 52621 52621 52621 52621 52621 53621 53221 533401 53461 53581 53581 53581 53581 53701 53701	ACGGTCGCGG GCCCGCGGA AGGCGTTCGA TCCTCGGCGGTCA AGTCGGTCGC CGCGGGTCAC AGTCGGCGGTCAC AGGCGGTCAC AGGCGGTCAC AGGCGGCCGT CCACCCCACC	CGGCTTCCTC GGCGCTGGGG GCACGCGGGG CACGAGGGGG CTCCGGGGAA CTCCGGGGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGGC CCAGGACGCC CGAACTCCTC CTCCTCCTTC CTCCTCCTTC CCCCCGCAG CCGCGCGGAC CCGCCGCAG CCCCCGCAG CCCCCCGCAG CCCCCGCAG CCCCCCGCAG CCCCCCGCAG CCCCCCCC	ACCGGGGGGGGAACGACGACGACGGCGGGGGGGGGGGG	COGGCTTOGA AGCAGOGCTT AGACGCTGCGA GCCTGGCGGCGA AGGCGGCGGCGA GCCTGGCGGCGGCGGCCGACACCCGACACCCGACACCCGACACCCCACCCCACCCCACCCCACCCCACCCCCACCCCCACCCC	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC CTTCTACGGT DTCTACGGT DGCGCTGCAC GGGGTCACG GGGGTCACG GGGGTCACG GGGGGAC CGAGGCACAG CTCGCAGGCC CGAAACGGA CTCGCACGTC CGAACGGCC CCACATCATC CGGACCGCTG CCGCCTGGCC ACTCGCCCGG CCTGAGCCCG CCTGAGCCCG CCTGAGCCCG CCTGAGCCCG CCTGAGCCCG CCTGAGCCCG CCTGAGCCCG CCTGAGCCCG CCTGAGCCCG CCACCCCCCAC	TTOGGCATCA ACCTCGTGGG ACCGGGGGTGT TACCGGAGGGTC CACCGGGGGC GTGATGGGCGGGGGGGGGG
40	52501 52501 52621 52621 52621 52621 52621 52621 53621 53261 53261 53261 53461 53521 535641 535641 535641 535641 535641	ACGGTCGCGG GCCCGCGGA AGGCGTTCGA TCCTCGGCGTCAT AGTCGGTCAT AGTCGGTCAT AGTCGGTCAT AGTCGGTCAT AGGCGGTCAA AGGCGTCAA AGGCGGTCAA AGGCGGTCAA AGGCGGCGTGT ACCCCGGCGGTGT ACCCCGGCCGTGT ACCCCGGCCGTGT ACCCCGGCCGTGT ACCCCGGCCGTGT ACCCCGGCCGTGT ACCCCGGCCGTGT ACCCCGGCCGTGT ACCCCGGCCGTGT ACCCCGGCCGTGT ACCCCGGCCGGCGTGT ACCCCGGCCGGCGTGT ACCCCGGCCGGTGT ACCCCGGCCGGCGTGT ACCCCGGCCGGCGTGT ACCCCGGCCGGCGTGT ACCCCGGCCGGCGTGT ACCCCGGCCGGCGGTGT ACCCCGGCCGGCGGTGT ACCCCGGCCGGCGGGCGGGCGGGGGGGGGG	CGGCTTCCTC GGCGCTGGGG GCACGCGGGC GTTCTTCCAG CACGAGGGGAA GTTCGGGGGAC GGAGGAGGGC CTCGGAGGGC CCAGGAGGGC GCGCGGGGGGGGGG	ACCGGGGGGGGGAACGACGACGACGGGGGGGGGGGGGG	COGGCTTOGA AGCAGOGGCT AGACGCTGCGA GCCTCGCGGA CCCTCGCGGA GCCTCGCGA GCCACCGGACT GCGCGGACT GCGCGACCGACC GCACCGACCG GCACCGACCG GCACCGACC GCACCGACC GCACCGACC GCACCGACC CACCGGACA CACCGACA CACCGCCACA CACCGCCACA CACCGCCACA CACCGCCACA CACCGCCACA CACCGCCACA CACCGCCACA CACCGCCACA CACCGCCACA CACCGCCACA CACCGCCACA CACCGCCACA CACCGCCCACA CACCGCCACACC CACCGCCACACCC CACCGCCACACCC CACCGCCACACCC CACCGCCACACCC CACCGCCACCCC CACCGCCACCCC CACCGCCCACCCC CACCCCCACCCCCACCCC CACCCCCCCC	CGCGGCGTTC DGCCCTGGAG GGGCAGTGAC CTTCTACGGT DTTCTACGGT DGCGCTGCAC GGGGCGCGCGC CGCGCGCGCGCC CGCGCGCGCC CGCGCCGC	TTOGGCATCA ACCTCGTGGG ACCGGGGGTGT TACCGCACCA CTGGAGGGTC CAGCCGGGC GTGATGGCTGCA GTCCTGATGG GTCCAACATCG GCCATGGAGG GCCATGGAGG GCCATGGAGG GCCATGGAGG GCCATGGAGG GCCATGCTGC GCCTTCCTCG GCCATGCTGC GCCATGCTGC GCCATGCTGC GCCATGCTGC GCCATGCTGC GCCACCCAGT AACGCCGGCC ACCGGGCC ACCGGGCC GACCACCTCG
40 45 50	52501 52561 52681 52681 52681 52681 52681 5381 53161 53281 53281 53581 53581 53581 53581 53761 53881 53881	ACGGTCGCGG GCCCGCGCGA AGGCGTTCGA TCCTCGGCGTCAA AGGCGTTCAC AGTCGCTCGCGTCAA AGGCGTTCGC TCGAGAAGCT TCGAGGACCCA AGGCGCGCGT ACCCCGGCGCGT ACCCCGGCCG ACGACACCC ACGAGACCCC TCGAGCACCC TCGAGCACCC ACGACCCC ACCCCCCCC	CGGCTTCCTC GGCGCTGGGG GCACGGGGG GTTCTTCCAG CACGAGGGGG GTTCTCGAGGAGGGG GTTCGAGGAGGGC CTCGAGGAGGGC CTCGAGGAGGGC CTCGAGGAGGGC GCGCAGGAGGGC GCCGCAGGAGGGC GCCGCAGGAGGGC GCCGCAGAGGGC GCCGCAGAGGGC GCCCGCAGAGGGC GCCCGCAGAGGGC GCCCGCAGAGGGCC GCCCGCAGAGGGCC GCCCGCAGAGGGCC GCCCGCAGAGGGCC GCCCGCAGAGGCC GCCCGCAGAGGCC GCCCGCAGAGGCC GCCCGCAGAGCCC GCCCGCAGAGCCC GCCCGCAGGCCC GCCCGCAGCCC GCCCCCCCC	ACCGGGGGGGGGGAACGGGGGGAACGGGGGGAACGGGGGG	COGGOTTOGA AGCAGOGOTT AGACGOTTGO TOGGOGGOTT GOTTOGGOGA AGGGGGGGGG GOTTOGGOGA GOCAGOGGGT GOTTOGGOGA TGOTTOGGOGA TGOTTGOGGG GOCAGOGGGG GOCAGGGGG GOCAGGGGG CACAGGTACA TCGCGGAGAC CACAGGTACA TCGCGGAGAC CACAGGTACA CACAGGTACA AAAGCACGC AAAGCACGC AAAGCACGC AAAGCACGC AAAGCACGC CCCACAGAC CCCACAGAC CCCACAGAC	CGCGGCGTTC CGCCCTGGAG GGGCAGTGAC CTTCTACGGT CTTCTACGGT CGCGCTGCAC CGCGCTCACC GGGGTCACC CGCGCGCGCC CGCGCGCCC CCGCGCCCC CCGCCCCC CCGCCCCCC	TTGGGCATCA ACCTCGTGGG ACCGGGGGTGT TACCGCACCAC CTGGAGGGTC CAGCCGGGGC GTGATGGCGGGT GCGCGGGGTT CCGCGGGGTT CCGCGGGGGGGG
40 45 50	52501 52561 52621 52681 52681 52681 52681 53681 53161 53281 53281 53581 53581 53581 53761 53881 53881 53881 53881	ACGGTCGGG GCCCGCGGA AGGCGTTCGA TCCTCGGGGTCAT AGTCGGTCAT AGTCGGTCAT AGTCGGTCAA AGGCGTTCAA AGGCGTCAA AGGCGTCAA AGGCGTCAA AGGCGTCAA AGGCGCGCGT CCGGGGGCGTGT ACCCCGGCGGTGT ACCCCGGCGGTGT ACCCCGGCCGT CCGGGCGCGT ACCCCGGCCGT ACCCCGGCCGT ACCCCGGCCGT ACCCCGGCCCG ACGACACCC ACGACACCC ACGACACCC ACGACACCC ACGACACCC ACGACACCC ACGACACCC ACGACACCC ACCCCACCCA	CGGCTTCCTC GGCGCTGGGG GCACGGGGG GTTCTTCGAG GACGAGGGGG GTTCGAGGAGGGG GTTCGAGGAGGGC GGAGGAGGGC GCGGGGGGGGGG	ACCGGGGGGGGGGAACGGGGGGAACGGGGGAACGGGGGAACGGGGGG	COGGOTTOGA AGCAGOGOTT AGACGOTTGO TOGGOGGOTT GOTTGGOGGA AGGGGGGGGGG	CGCGGCGTTC CGCCCTGGAG GGGCAGTGAC CTTCTACGGT CTTCTACGGT CGCGCTGCAC CGCGCTCACC CGCGCGCAC CGCGCCCCC CCGCCCCCCCCCC	TTGGGCATCA ACCTCGTGGG ACCGGGGGT TACCGCGGGGC CTGGAGGGTC CAGCCGGGC GTGATGGCGGGT GCCGGGGGT GCCGGGGGGGGGG
40 45 50	52501 52561 52621 52681 52681 52681 52681 53641 53161 53261 533401 53581 53581 53581 53761 53881 53881 53881 53881 53881	ACGGTCGGG GCCCGGGGA AGGCGTTCGA TCCTCGGGGTCAC AGTCGGTCAC AGTCGGTCAC AGTCGGTCAC AGTCGGGGTCAC AGGGGGTGAT TCGAGGACAC CCACCCCAC ACGGGGGGGTGT ACCCCGGGGGCGGTGT ACCCCGGGCGGTGT ACCCCGGGCCGGTGT ACCCCGGGCCGGTGT ACCCCGGGCCGGTGT ACCCCGGCCGGTGT ACCCCGGCGGTGT ACCCCGGCGGTGT ACCCCGGCGGTGT ACCCCGGCGGTGT ACCCCGGCGTGT ACCCCGGCGGTGT ACCCCGGCGGTGT ACCCCGGCGTGT ACCCCGGCGTGT ACCCCGGCGTGT ACCCCGGCGTGT ACCCCGGCGTGT ACCCCGGCGTGT ACCCCGGGGGTGT ACCCCGGGGGTGT ACCCCGGGGGTGT ACCCCGGGGGGTGT ACCCCGGGGGGGTGT ACCCCGGGGGGGGGG	CGGCTTCCTC GGCGCTGGCG GCACGCGGGC GTTCTTCCAG CACGAGGGGAA GTTCGGGGAA GTTCGGGGAC GCAGGACGGC CCAGGACGGC CCAGGACGGC GCAGGACGGC GCAGGACGGC GCAGGACGGC GCAGGACGGC GCAGCCGCAC CCGAACTCCTC CTCCTCCTTC GCCCCCGCAG CACCCCGCAG CACCCCGCAG CACCCCGCAC CGACCTCCTC GCCCCCGCAC CACCCCGCAC CGCCCGCAC CGCCCGCAC CGCCCCGCAC CGCCCCGCAC CACCCCCGCAC CGCCCCGCAC CGCCCCGCAC CGCCCCGCAC CGCCCCGCAC CGCCCCGCAC CGCCCCCCCC	ACCGGGGGGGGGAACGGGGGGAACGGGGGAACGGGGGAACGGGGGG	COGGOTTOGA AGCAGOGOTT AGACGOTTGO TOGGOGGOTTGO TOGGOGGOTTGO TOGGOGGOTT COCTGOTGOTGO AGGGOGGOTT GOTTGOGGOGA GOGGOGGOTG TGCTGOTGGO GOGGOGGOTG GOGCGAGAG GOGCGAGAG CACAGGTACA TCGCGGAGAG CACAGGTACA TCGCGGAGAG CACAGGTACA TCGCGGAGAG CACAGGTACA TCGCGCAGAG TCGCGCAGAG CACAGGTACA AAAGCACGCT AAGCGCAGAG TCGGCGAGAC TCGGCGAGAC TCGGCGAGAC TCGGCGAGAC AAGCACGCT AAGCGAGACC CCCACCAGAC TCGGCGAGAC	CGCGGCGTTC CGCCCTGGAG GGGCAGTGAC CTTCTACGGT CTTCTACGGT CGCCTGGAC CGCGCTGAC CGCGCGGAC CGCGCGGAC CCACATCATC CGCACCGCGC CCACATCATC CGCACCGCGC CCACCCCGCCC CCACCCCCCCCCC	TTGGGCATCA ACCTCGTGGG ACCGGGGGT TACCGCAGGGC CTGGAGGGTC CAGCCGGGC GTGATGGCT GCCCGGGGT GCCCGGGGT GCCCGGGGGGGGGG
40 45 50	52501 52561 52681 52681 52681 52681 52681 53681 53161 53281 53281 53281 53581 53581 53581 53581 53761 53881 53641 53881 53881 53881 53881 53881	ACGGTCGGG GCCCGCGGA AGGCGTTCGA TCCTCGGGGTCAT AGTCGGTCAT AGTCGGTCAT AGTCGGTCAA AGGCGTTCGA AGGCGTCAA AGGCGTCAA AGGCGTCAA AGGCGTCAA AGGCGCCCA CCACCCAA AGGCGCCCA CCACCCAA AGGCGCCCT CCGGCGCCCT ACCCCACCCA ACGCCCCCACCCA ACGCCCCCCACCCCACCCCACCCCCACCCCCACCCCCACCCCACCCC	CGGCTTCCTC GGCGCTGGGG GCACGCGGGC GTTCTTCCAG CACGAGGGGAA GTTCGGGGGAA GTTCGGGAGGGC GGAAGGGGC GGCGGGGGGGGGG	ACCGGGGGGGGGGAACGGGGGGAACGGGGGAACGGGGGGG	COGGOTTOGA AGCAGOGOTT AGACGOTTGO TOGGOGGOTT GOTTOGGOGA GOTTOGGOGA GOTTOGGOGA GOGAGOGGOTT GOGTGATCAC GOGAGOGGAT TGOTGOTGGO GOGACGAGOT GOGAGGAGA CACAGGTACA CACAGGTACA CACAGGTACA TOGGGGAGAC CACAGGTACA TOGGGGAGAC CACAGGTACA CACAGGTACA CACAGGTACA CACAGGTACA TOGGGGAGAC TOGGGGACT AAGGGTGGTO ACGGGGGACT ACGGGGGACT CACAGGTACACA CACAGGTACA CACAGGTACA CACAGGTACA CACAGGTACA CACAGGTACA CACAGGTACA CACAGGTACACACACACA CACAGGTACACACACACACACACACACACACACACACACA	CGCGGCGTTC CGCCCTGGAG GGGCAGTGAC CTTCTACGGT CTTCTACGGT CGCGCTGCAC CGCGCGCGCC CCACACACAC CCACACACAC CCACACACA	TTOGGCATCA ACCTCGTGGG ACCGGGGGT TACCGCGGGGC CTGGAGGGTC CAGCCGGGCC GTGATGGCGGGT GCGCGGGGTT CCGCGGGGTT CCGCGGGGGGGG
40 45 50	52501 52561 52681 52681 52681 52681 52681 53681 53161 53281 53281 53281 53581 53581 53581 53581 53761 53881 53641 53881 53881 53881 53881 53881	ACGGTCGGG GCCCGCGGA AGGCGTTCGA TCCTCGGGGTCAT AGTCGGTCAT AGTCGGTCAT AGTCGGTCAA AGGCGTTCGA AGGCGTCAA AGGCGTCAA AGGCGTCAA AGGCGTCAA AGGCGCCCA CCACCCAA AGGCGCCCA CCACCCAA AGGCGCCCT CCGGCGCCCT ACCCCACCCA ACGCCCCCACCCA ACGCCCCCCACCCCACCCCACCCCCACCCCCACCCCCACCCCACCCC	CGGCTTCCTC GGCGCTGGGG GCACGCGGGC GTTCTTCCAG CACGAGGGGAA GTTCGGGGGAA GTTCGGGAGGGC GGAAGGGGC GGCGGGGGGGGGG	ACCGGGGGGGGGGAACGGGGGGAACGGGGGAACGGGGGAACGGGGGG	COGGOTTOGA AGCAGOGOTT AGACGOTTGO TOGGOGGOTT GOTTOGGOGA GOTTOGGOGA GOTTOGGOGA GOGAGOGGOTT GOGTGATCAC GOGAGOGGAT TGOTGOTGGO GOGACGAGOT GOGAGGAGA CACAGGTACA CACAGGTACA CACAGGTACA TOGGGGAGAC CACAGGTACA TOGGGGAGAC CACAGGTACA CACAGGTACA CACAGGTACA CACAGGTACA TOGGGGAGAC TOGGGGACT AAGGGTGGTO ACGGGGGACT ACGGGGGACT CACAGGTACACA CACAGGTACA CACAGGTACA CACAGGTACA CACAGGTACA CACAGGTACA CACAGGTACA CACAGGTACACACACACA CACAGGTACACACACACACACACACACACACACACACACA	CGCGGCGTTC CGCCCTGGAG GGGCAGTGAC CTTCTACGGT CTTCTACGGT CGCGCTGCAC CGCGCGCGCC CCACACACAC CCACACACAC CCACACACA	TTOGGCATCA ACCTCGTGGG ACCGGGGGT TACCGCGGGGC CTGGAGGGTC CAGCCGGGCC GTGATGGCGGGT GCGCGGGGTT CCGCGGGGTT CCGCGGGGGGGG

	1441		:33393h33h3	38832 3 3782	7,735,737,37	0000000000	-030ATOOA01
	11:11:1		3873037080	30033300807		90%300%0T0	3703355333
	14361		12000000000	2000	ATTRATTOTA	25000000000	COCCCCCACO
	4 7						08.0000000A0A
5	14:1	1004004004	0000000000	3,27,20,00	ABACCTONT	72.002.35A0	2727232222
	- 154	7037037307	00770223333	CAGACCGCTA	10000034038	307309999	OTGOMOROOS
	54671	0000000000	3270010370	000000000000000000000000000000000000000	00870088010	GROGOTO	0700000000
	54661	- Garana an	0000100000	anak mame ma	000000000000	2AA33A27A3	mccamcaca
	1956.		-200001M002	100000000000	300A33A00A	33T3G0G0A0	
1()				ACCECCOCC	~	37.3303000 37.33030000	
119	1 T 1 T 2			000400000		- 371	0131030736
	34641		GTGGGTCGGC	อกเอเก็กเป็นโลเล		COTSCTOOMS	0000000000
	14901	007700703A	ACTOGOGGCG	0303003609	MUUNGGI ÜÜG	empogacong	CTGCACGAAC
	54961	TOGTORTOGR	GACGCCGCTC	GTGCTGCCCG	00400000000	1919903610	o.canon
	35001	10000035500	CGACGACACG	3630630339	DESTOROSST	22A232GCGG	GOCGACGGCT
15	£5081	13330071913	GACCCGACAC	GCCGGCGGA:	100T0GGCAC	GECACOGGCA	DUGGGUAUGG
	55141	00A066A000	GGCACCCTGG	20300000033	ANGCCGGROO	GBTCSACGTC	GCCGACGTCT
	50000	REGGRESE	CGAGGACATC	GGGTACTCCT	MOGGACCGGG	CTTCCGGGGG	CTGCGGGGCG
	35261	0073603000	COGCGACACC	GTGTACGCCG	REGETTGCGTT	COCCGACGAG	CAGAGCGCCG
	55321	ACCCCCCCC	TTTCACGCTG	CACCCGGGG	TGCTCGACGC	DECETTOCAG	G003606060
20	86381	1000000001	CGACGCACCC	GGGGGGGGG	noogactgoo	3770703770	CAGGACGTCC
~ ~		2212001000		ACGCGGCTGC	SGSTOACSST	253223CSA2	GGCGAGCGCA
	3050		000000000000 0ATGA00000	0000409301	A001031300	0000000000	2003730707
		3.5000000000			3007307303	COUGSTOTES	ACCGAGCTCC
	EBEET	030300037A	OG TGGNAGGO	TCCGGTGACG	3037037337	001033.0100	GACCCCGGCC
2.5	35601		cccaracass	SACGATOSSC			
25	1,5,6,5,1	A00003A0GT	TCCGGCGGGGC	ACCCGGGAGC	TORCOGOGG	CSTCCTCGGC	GCSCTCCAGC
	35741	GCCACCTCTC	CGCCGCCGAG	GACACCACCT	TGGTGGTACG	GACCGGCACC	GGCCGGGCCG
	55801	073003333	CGCGGGTCTG	GTCCGCTCGG	CGCAGGCGGA	GAACCCCGGC	CGCGTCGTGC
	55661	TOGTOGAGGO		ACCTCGGTGG		CGCGTGCGCC	GCGCTGGACG
	35921	AACCGCAGCT	GGCCGTCCGG		TOTTOGOGGC	GCGGCTGGTC	CGGATGTCCG
30	55981	ACCCCGCGCA	CGGCCCGCTG	TCCCTGCCGG	ACGGCGACTG	GCTGCTCACC	CGGTCCGCCT
	56041	CCGGCACGTT	GCACGACGTC	GCGCTCATAG	CCGACGACAC	GCCCCGGCGG	GCGCTCGAAG
	5€101	CCGGCGAGGT	CCGCATCGAC	GTCCGCGCGG	CCGGACTGAA	CTTCCGCGAT	GTGCTGATCG
	56161	COCTOGGGAC		GCCACGGCCA		GGCCGCGGGC	GTCGTGGTGG
	86221	AGACCGGGCC			COGGOGACCG	GGTGTTCGGC	CTGACCCGGG
35	56281	GCGGCATCGG		GTCACCGACC	33CGCTGGCT	GGCCCGGATC	CCCGACGGCT
27.27	56341	GGAGGTTCAG	CACGGCGGCG	TCCGTCCCGA	TOSTGTTOGO	GACCGCGTGG	TACGGCCTGG
			CACACTGCGC	GCCGGCGAGA	ASGTCCTCCT	COACGOGGCC	ACCGGCGGTG
	56401	TOGACCTOGG		ATOGODOGOO	ACCTGGGGGG	COASSTORAC	GCCACCGCCA
*	56461	TOGGCATGGC	CGCCGCACAG		********	CGACACGCAC	ATCGCCGACT
473	15121	ARCOCOLATE	GCASCACGTC	CTGCGCGCCC	CGCGCATGGA	CUTCUTCUTG	AACGCGCTGA
40	3 45 8 1	CTCGGACGAC		ACCCCTTTCC			TTCGTCGAGA
	56641	000303A3TT	CATCGACGCG	TOGOTOGACO	TGCTGGACGC	CGACGGCCGG	
	56701	TGGGCCGCAC		GACCCGGCCG	CGATCSTCCC		CCGTTCGACC
	56761	TGCTGGACGC	GGGCGCCGAC	CGCATCGGCG	AGATUULGGG	CGAACTGCTC	0000101.00
	5.6831	A000333000	GCTGSAGCCG	CTGCCGGTCC	GTGCCTGGGA	CGTCCGGCAG	01.A01011A10
45	5,6881	CGCTCGGCTG	GATGAGCCGC	GCCCGCCACA	TOGGCAAGAA	CGTCCTGACG	CIRCCCCCCCC
	86941	ogorochico	GGAGGGGGCC	GTCGTCCTCA	CCGGCGGCTC	CGGCACGCTC	GCCGGCATCC
	97361	0000000000	CCTGCGCGAA	CGGCATGTCT	ACCTGCTGTC	CCGGACGGCA	CCGCCCGAGG
	50061	ggangenegg	CGTCCACCTG	CCCTGCGACG	TOGGTGACCG	GGACCAGCTG	GCGGCGGCCC
	20101	magagaaaaa	BODDBDDAADA	ATCACCGCCG	TGGTGCACCT	CGCCGGTGCG	CTGGACGACG
50	57181	SCACCGTOGO	GTGGGTGAGG	CCCGAGCGTT	TCGACACGGT	GCTGCGCCCG	AAGGCCGACG
	57241	ATERMOOPER	CCTGCACGAG	CTGACGAAGG	AGCAGGACCT	CGCCGCGTTC	GTGCTCTACT
	555-	0000010011	CORDOTTECTC	GGCAACGCCG	GCCAGGGCAA	CTACGTCGCC	GCGAACGCGT
	5 7 3 5 1	201633223	COGCOTOCIC	CTRCGCCACG	gttengagge	GCCGGCCCTC	TOGATOGOGT
	5.301	accessment	GCTCGCCGWG	CIDCOCOCCACI	00000000000	CGGCGAAGCC	GACCGGGACC
22	D 9		DI OUNDUNIO	AGCGGGGTCA	02000007377	CECCLTRCTLC	7767206230
55	5/451	-GGATGCGGCG	CAGCGGTTTC	COGOCCATCA	#000000MM0H	GGGCATGCAC	BOACOACA
	57541	CGGCCGGCCG	CACCGGAAGT	CCCGTGGTGG	70000000000	GCTCGACGAC	COCCOCCACA
	57601	raccacracr	SCGCGGCCTG	UGGUGGACGA		GGCCGCCGTC	
	57661	CETTOCGCCGA	CUGGCTCGCC	GCGCTGACCG		CGCCGAAGCG	7 mode00000
	27771	Φ_{i}^{-1}	CIRCICORCE	-ACCATACTES	GCCACGTGGG	TGGCGAGGAU	E TUUUUUU
60	57781	CGGCGGCGTT	CAAGGACCTC	GGCATCGACT	GGGTGAGGGG	GGTCCAGCTG	CGCAMUBCCC

					33		
	-41	7080008300		133773W.03	11740330731		DOGROOCOGE
		A037307000	COCCARGOTT	9903N03F4N0	73573333A	1030306000	GTCGTGCCCC
	* 11 M.	-03A0030331	080000000	3030A03A03	A000000000	GATTGTTGGGA	ATGGCCTGCC
			7333372533	7280023863		77770773378	70003304000
5	5.50	A0000AT0A0	3383TT0003	A000H10303	007333A701	00803068T0	TARBACCOOS
	5 4 1 4 1	A00003A030	BATOBSCNAG	A007720702	GGCACGGT03		DBCBCGACAB
	55111	3000030	SSECTOTIC	330AT 0A300	000000000000000000000000000000000000000	COTOGOGATO	GACCCGCAGC
	5+241	8000337307	CCTGGAGACG	T00T333A33	0317032440	-0900990ATC	ACCCCGGACT
	56321	1340003123	CMGCGACACC	3303T3T723	70000000077	TECOMACCON	TRIGGGRACES
10	1331	GTGGGGACAC	CGAGGGGTTC	GGGGGGAGGG	SOTOSOASAS	CASTSTSCTS	Todagacada
	Fr441	TOTOGTACTT	CTACGGTCTG	GAGGGTCCGG	COGTONOSST	DGACACGGCG	TETTCGTCGT
	5.4555	ogdTggTggC	GCTGCACCAG	GCCGGGCAGT	CGCTGCGCTC	CGGCGAATGC	Tegeregeed
	88881	TGGTCGGCGG	CGTCACGGTG	Argadorord	ccagcaact?		TOCCGGCAGC
	88621	3066001060	GUCGGACGGC	CGGGCGAAGG	C3TT0G00000	GBGTGCGGAC	GGCACGAGCT
15	55681	T09009A933	TGCCGGTGTG	CTGATOGTOG	50533072T2	GGACGCCGAA	CGCL=CGGTC
	58741	ACACCGTCCT	GGCGGTCGTC	CGTGGTTCGT	OGGTCAR 20%.	03AT 33T G00	TOCAACGGGC
	88801	Tamoggagaa	GAACGGGCCG	TOGGAGGALL	GGGTGATGCG	BONGGCCCTG	GCCAACGCCG
	56861	GGGTCACCCC	GGCGGACGTG	GACGCCGTCG	AGGCCCACGG		AGGCTGGGCG
	58921	ACCOCATOGA		GTACTGGCCA	COTACGGACA	GGAGCGCGCC	LOCACACATA
20				AACATGGGCC	ACGCCCACCC	030GT00GG0	STOSCOGGCA
20	58981	TGCTGGGGTC	GCTGAAGTCC	68866663336		0000A00000	CACGCCGGCA
	19041	TCATCAAGAT	GOTGCAGGCC	- Waldestanes	933A30T000		
	59101	AGCOGTCGCC	GCACGTCGAC	TGGACGGCCG	000000000000000000000000000000000000000	TTCGTTCGGG	7733777737
	E9161	demagadadea	GACCGACCGG	CCACGGGGTG			GTGAGCGGCA
3.5	59021	CCAACGCCCA.		GAGGCTGGAC	COSTAACGOA	GACGCCCCCC	GCATCGCCTT
25	59281	COGGTGACCT		GTGTCGGCAC	GUTUMULGEN	AGCCCTCGAC	GAGCAGATCC
	59341	GCCGACTGCG		GACACCACCC		COGGGTGGCC	GTGGCACAGA
		CGCTGGCCCG		TTCGCCCACC	GCGCCCTGCT	GOTOGGTGAC	
		CCACACCCCC		CCCGACGAAC	TOSTOTTOST	CTACTCCGGC	CAGGGCACCC
			GATGGGCGAG		CCGCCCATCC		GACGCCTGGC
30	59581	ATGAAGCGCT	CCGCCGCCTT	GACAACCCCG	ACCCCCACGA	CCCCACGCAC	
		TGCTCTTCGC				GTCCTGGGGC	
	59701	ACGCGGTCAT	CGGCCACTCG	CTGGGCGAGA	TCACCGCGGC	GCACGCCGCC	
		CGCTGGACGA				CCTCATGCAC	
		CACCCGGTGC				SGCACGCCAG	
35		CGGGCGTGGA				CGTGCTGTCC	GGGGACGAGG
	59941	ACGCCGTGCT				dogcorgood	SCCCCGCACG
	60001	CCGGGGCACTC		GAGCCCGTGG		SCTCSCCACC	
	60061	TOOGCTACCA	CCCTCCCCAC	ACCTCCATTC		CACCASCGCT	GAGTACTGGG
	60121	CCGAGCAGGT	CCGCAAGCCC	GTGGTGTTCC		GORGORGTAC	CCGGACGCCG
40	60181	TGTTCGTGGA	GATOGGCCCC	GCCCAGGACC		COTTOGROGGG	
	60241	AGAACGGCAT	CGCGGGACGAG	GTGCACGCGC	TGCACACCGC	GCTCGCGCAC	
	60301	GCGGTGCCAC	SCHOGACTGG	CCCCGCATCC	TOGGGGGTTGG	GTCACGGCAC	GACGCGGATG
	60361	TGCCCGCGTA	CGCGTTCCAA	CGGCGGCACT	ACTGGATCGA	GTCGGCACGC	COGGCCGCAT
	60421	CCGACGCGGG	CCACCCCGTG	CTGGGCTCCG	GTATCGCCCT	CGCCGGGTCG	ccagaccaga
45	66481	TGTTCACGGG	TTCCGTGCCG	ACCGGTGCGG	ACCGCGCGGT	STTCGTCGCC	SAGCTGGCGC
, ,	80533	TGGCCGCCGC	GGACGCGGTC	GACTGCGCCA	CGGTCGAGCG	GCTCGACATC	GCCTCCGTGC
	80801	0000000000	GGGCCATGGC	CGGACGACCG	TACAGACCTG	GGTCGACGAG	COGGCGGACG
	60661	ACGGCCGGC3	CCGGTTCACC	GTGCACACCC	GCACCGGCGA	CGCCCCGTGG	ACGCTGCACG
	20001 20001	CCGAGGGGGT	CCTCCCCCC	CATGGCACGG	CCCTGCCCGA	TGCGGCCGAC	GCCGAGTGGC
50	20727	CCCCACCGG3	CGCGGTGCCC	GCGGACGGGC	TGCCGGGTGT	GTGGCGCCGG	GGGGACCAGG
2,0	60.01 60241	TCTTCGCCGA	GCCCGLGGTG	GACGGACCGG	ACGGTTTCGT	GGTGCACCCC	SACCTGCTCG
	20001	ACGCGGTCTT	ouccondito	GCCGACGGA	GOOGGOAGOO	GGCCGGATGG	CSCGACCTGA
	200501 200501	CGGTGCACGC		100011001A	GOGCCTGCCT	CACCCGGCGC	ACCGACGGAG
	00901	CCATGGGATT	GIGGGGGGGG	Treedingide	GCCCTGCCGGT	3.100000000 3.770370600	GEGGCGGTGA
5.5	01021	CCATGGGATT CGCTGCGGGA	1000000011C	COSTCCCCC	CCGAGGAGT	TELECTORICE OF THE STATE OF THE	CACCGGTTGG
55	51081	AGTGGCTCGC	acmacasas s	CCCTCCGGCT	ACCOMOUNT.	- dabagarese.	CATGTCCTGA
	61141	AGTGGCTCGC TOACCGCCGC	OGTCGCCGAG	CACCCCCATAC	POSSIONALI		10000000A
	01201	TOACCGCCCC	TURN DUCGAC	GALULTOARD	MCMIANCIAC	- 0.500000000000000000000000000000000000	
	01261	COCCOCCT	GREUBUCCTG	CORRURADO	manageneers	OT DESCRIPTO	onecleated
7.0	61321	ACACCACCAC	CGACCCCCCC	GGCGCCACCG	10,500,000	CATCACCOCACC	CTCCNGAACG
60	61381	AACACCCCCA	CCGCATCCGC	CTCATCGAAA	COMPLACIO	COMUMBUCCE	0,00000100

		0000000000	0.000T00T0	CRCCCCARC	~~~~~~~	201201222	000000000
	5 - 2 11 1			* agt got aga	21020110110	DUMICACATOR C	1133ACCACC
	- 1		COCCCTCCAU	ACCACCACCC	ema carredra	inconcora	remobilitional
	81581	ACCCCATCAT	CATCACCIGG	GGCTCCGGCA			CGCCACCTGA
	2000	ACCAPCICAL	15000550000	2727223333	202222222	COACCETAGE	CCCGGCACCC
5		1.0000000000		A * A A A A A A A A A A	ANCTIGOCAL	21.30000100	CACATOCCCC
-'	11.51			0.01000000	0-10-00-00		CTCCACGCCC
	0 1 4 1	AACCCCCTAAC		_MUM	SCAUCURCAN	2270232072	
	81691	TORCOCCOR		ACCUTTOUT	a con cara a constant of the c		TGGCACCTGC
	31881	ACCACCTOAC	COMMECCAN	CCCCTCACCC	ACTTCGTCCT	CTACTICAGI	GCCGCCGCCG
	61921	TOOTOGGCAG	CCCCGGACAA	GGNAACTACS	CCGCCGCCAA	20000000000	GREGEETEG
10	61961	202200000	CCACACCCTC	3600AA0000	CONCOTOCAT		ATGTGGCACA
10				CAACTOGAOG		RGACCGCATC	CGCCGCGGCG
	61041	CCACCAGGAC	CCTCACCGGA				
	62101	GTTTCCTCCC	GATCACGGAC	GACGASSSCA		DBAGGCGGCC	GTCGGCTCCG
	62161	GCGAGGACTT	CGTCATGGCC	GCCGCGATGG	ACCOGGCACA	GOOGATGACC	GGCTCCGTAC
	62021	CGCCCATTCT	GAGOGGCCTG	CGCAGGAGCG	CGCGGGGGGT	CGCCCGTGCC	GGGCAGACGT
15	6281	TOGOCCAGOG	GC PCGCCGAC		COGREGGES	ngeggegete	ACCACCCTCG
10			CACGGCCGCC	GTGCTCGGCC			GCGCCGACCA
	62341	TOT JACGO		3100103300	mon coccos		AACCGGCTCG
	62401	CGACGTTCAA	GGACCTCGGC	ATCGACTOGO	TCACCGCGA.		
	62461	CGGAGGCGAC	CGGGCTGCGG	CTGAGTGCCA	000106101.		ACACCTCGGG
	62521	TOCTOGOCGO		ACCGATCTGT	TOGGCACGGC	COTGCCCACG	CCCGCGCGGA
20	62581	CGGCACGGAC	CCACCACGAC	GAGCCACTCG	CGATCGTOGG	CATGGCGTGC	CGACTGCCCG
~ 0	62641	GCGGGGTCGC	CTCCCCCGAG	GACCTGTGGC	AGCTCGTGGC	STOCGGCACC	GACGCGATCA
	02041		ORCCOCCOCC	GGCTGGGACA	TOGACCGGCT		GACCCGGACG
	CL UI				3077007030		GGCTTCGATG
	62761	CCCCCGGCAA	GACCTACGTC	Capamagaaa			
	62621	CCGCGTTCTT	CGGCATCAGC	CCGCGCGAGG	CACGGGCCAT		CAGCGCGTCA
25	62881	TOOTOGARAC	CTCCTGGGAG	GCGTTCGAGA			ACGCTGCGCG
	62941	GCAGCGACAC			TOTOCCATGG	STACGGCGCC	GGCGTCGACC
	63001	TGGGCGGGTT	CGGCGCCACC	GCCACGCAGA	ACAGCGTGCT	orcoggaagg	TTGTCGTACT
		1000000011	GGAGGGCCCG		TOGACACCGC		TOGOTGGTCG
	63061	TOTTOGGOAT	COMBIGUICE	GCCG.CACCG			CTCGCCGGCG
	63121	CCCTGCACCA	GGCGGCACAG	GCGCTGCGGA	CTGGAGAATG		
30	63181	GTGTCACGGT	GATGCCCACC	CCGCTGGGCT	ACGTCGAGTT		CGGGGACTCG
	63241	CCCCCGACGG	CCGTTGCCAG	GCCTTCGCGG	AAGGCGCCGA		TTCTCGGAGG
		GCGCCGGCGT	TCTTGTGCTG	GAGCGGCTCT	CCGACGCCGA	GCGCAACGGA	CACACCGTCC
	63361	TCGCGGTCGT	CCTCTCCTCC	GCCGTCAACC		CTCCAACGGC	ATCTCCGCAC
		CCAACGGCCC			SCCAGGCCCT	CGACAAGGCC	GGGCTCGCCC
			CICCCAGCAG	CGCGTCTTCC			GACCCGATCG
35	63461	CCGCCGACGT	GGACGTGGTG	GAGGCCCACS	GCACCGGAAC		
	63541	AGGCACAGGC	CATCATCGCG		AGGACCGCGA		TACCTCGGTT
	63601	CGGTCAAGTC	GAACATCGGA	CACACCCAGA	CCACCGCCGG	TGTCGCCGGC	GTCATCAAGA
	63661	TGGTCATGGC			CGAAGACACT	GCACGTGGAC	GAGCCGTCGT
•	55701	CGCATGTGGA	CTCCACCGLG	GGTGCGGTGG	AACTGCTCAC	CGAGGCGAGG	CCGTGGCCCG
10			CCCGCGCCCGC	CCCCCCCTCT	CGTCGCTCGG		ACGAIACGCCC
40		ACGCGGGACG		0.00000000		GCCGTCTGTT	• • • • • • •
	63841	ACGTGATCCT	TGAGGGTGTT	CCCGGGCCGT	COCOTO LOSA	2000121311	CACCGCCTCC
	63901	TGCCGTTGCC	GGTGTCGGCT	CGGAGTGAGG	CGAGTCTGCG	ם ג' שטמטיטיטיטיטיטיטיטיטיטיטיטיטיטיטיטיטיטיט	GAGCGGCIGG
	22421	DOCGOTATOT	CCCCCGGGAGT	-GTGGATGTGG	CCGCGGTCGC		GIGGGIGAGC
	2:001	CTCCTCTCTT	COGTCACCGT	GCGGTACTGC	TGGGTGATGC	CCGGGTGATG	GGTGTGGCGG
4.5	24001	TGGATCAGCC	CCCTACCCTC	THE CHECKENTY	CCGGGCAGGG	TGCTCAGTGG	GTGGGCATGG
45	54061	GTGTGGAGTT	GLODDARIOUS GLODGARIOCO	mcmccccmca	TOCOCCOCCOC	mimogaggag	TGTGCGCGGG
	64141	GTGTGGAGTT	GATGGACUGI	10.0000.01	CCCACAMOMM	200000000000000000000000000000000000000	CTTCTCCCC
	64201	CGTTGTTGCC	GCACACGGGC	TGGGATGISU	1101AuAucc	33030333003	GAIGIGGGG
	64261	AGCGGGTGGA	GGTGGTCCAG	CCGGCCAGCT	GGGCGGTCGC	a a la Challe i e	GC JGCMC1G1
	21331	COCACCCCA	COSCOTOGTA	CCCGACGCGG	TGATCGGACA	CTCCCAGGGC	GAGATUGUGG
50	61321	CGGCGTGCGT	ageneggggg.	CTCAGCCTTG	AGGACGCCGC	CCGCGTGGTG	GCCTTGUGUA
2.0	74771	GCCAGGTCAT	220000000	CTGGCCGGGC	GGGGAGCGAT	GGCTTCGGTG	GCATTGCCGG
	04441	CCGGTGAGGT	- acmamacma	CTGCCCCTCT	devade de const	3037110660	CCCGCCTCGA
	64501	CCGGTGAGGT	CGGTCTGGTC	GNGGGGGTGT	* OCT COUCCE	07.000000000	Characterine
	64561	CAGTCGTGGC	CGGCGAGCCG	TCGGCGGTGG	AGUACG POLL	THI DUDUUME	-maniford
	51501	CCCTCCCTCC	CCCTCGTATC	GCCGTCGLCT	ACGCCTCCCA	CACGCCCCAC	GIGGMAGGGA
55	20201	TOGROOMOGA	TOTCGCTGAG.	-GTACTGAAGG	GAGTTGCAGG		100010001
الدال	0400T	GGTGGTCGAC	0.0100010.10	SCCTSGGTGA	CCGAGCCGGT	GGATGAGAGT	TACTGGTACC
	04/41	GGTGGTCGAC	DEACADOLEC	0001000100	CARCARTOR	GGAGCTGGAC	GGGTCCGTGT
	€4801	GGAACCTGCG	TEGECCEGTC	GUGGLGGMCG	000000000000	-03M00133M0	- CD CD C C C C C C C C C C C C C C C C
	64861	TOGTGGAGTG	CASCGCCCAT	CCGGTGCTGC			CHURL30103
	21001	COMCOMMOCO	_ <u> </u>	CGCGGCTGGG	- Alabba Albaba -	- ひれしひひじには エン	
60	64981	GGACCCTGGG	CGCGGCAGTG	GACTGGGACA	CGGTGGTCGA	ACCGGTGCCA	GGGCGGCTGC

				D7001-0-0	0.200222222	30000000	
		TCGATCTSCC			GGUGUTATI A	30033AA303	GCCGGTGCCA
	45101	COGRECATE	-03033000330	CTGACACGCC	CASCASATOR	CATGOTGGCC	GCCATCACGC
	45141	CACTACCCC	CGACGACGGT	GGTGTTGTTC	70%0000000	GATOTOGITE	CGCACGCATC
	11.	2002200253		31203333331	2007200000	2000000A00	300000000000000000000000000000000000000
5		nacration.	10100000000	GACGAGAGGA		ASTSSATSAR	
	43341	- ಮಕ್ಕಾನಕ್ಕವಾಗಿತ - ಕರಕಾರ್ಯಕ್ಕವಾಗಿತ		- 303 % 0233 <i>h</i> 3	000000000000	0703375A30	0.33108106
		707.00.000		00.00000000	mag. 2000200	3:400:-000	J. GERMOOMS
	2141	iliahush330	Judhludildu	Canaliances		- U = 0 3 LT4/3/9/	
	db461	GGACOCGGCA	GGCCAGCGGC	ACCCTGACCC		COMPACCIOCO	
	EEEC:	GTGTTGTTGG	TGCGGARGCCG	TTCTTGGCAGT		CACTGCCGCG	
10	35511	COTTOGGAGTT	CTROTTGGGG	CTGGACGCC	TG33301A003	TTTCGGACCC	ATGTTCCGCG
	68641	GAMTG0G330		GATGGTGAGA	COCTOTACE	1343370300	CTCCCCGAGG
	, , , , , ,	2,0000000000		GRETTORICE		purgordaka	3234464404
	33-35		CGACGTGGAC			CARCOTOCAA	
			CCTSCTCATG	0.03001000	- Adda a da d		
1.5	68821	COTGGCACGG	CGTCCGGTTC	CACGOGACGG	GOGGGGGGG		GOGGTCGTAC
15	45551	2.7	0330010033	CTGCATGCCG	030ACA9133	COSTEDORAGE	
	85941	TOGALUGOT		TCCCCGGAAG	DGGACULUUU	GCCCGCCGAT	
	86001	J3GTCGGGTG	GGCCCCGGTG	CCCGTACCTS	00003300363	TOOGTOOGAC	GCGGACGTGC
	55051	TGACGCTGCG	CGGCGACGAC	GCCGACCCGC	TOGGGGAGAG	COGGGACCTG	ACCACCCGTG
	66121			GCCGACCGGC		COAGGTGACC	GSTGGCCTCG
20	66181			CTGGTCCGCA		CGAGCAGCCC	
±()		TOGUCAAGGU	PROCESS ACCE	GGAGAGGTCC			
	66241						
	86301	CACTCGGCGA	GCCCCATGTS	CGGCTGCGCG	ACGGCCTCTT	CGAGGCAGCC	CGGCTGATGC
	56361	GGGCCACGCC	GTCCCTGACG	CTCCCGGGACA	COGGGTCGTS	GCAGCTGCGG	COGTOCGCCA
	66421	COGGTTCCCT	CGACGACCTT	GCCGTCGTCC	CONCOGRACIO	COCGGACCGS	CCGCTCGCGG
25	66481	COGGCGAGGT	GOGGATCGCG	GTACGCGCGG	CGGGGCCTGAA	CTTCCGGGGAT	STCACGGTCG
	22531			GCGCGTCCGC		33CCGCGGGT	GTCGTCCTGG
	66601	AGACCGGCCC	00.00000000	6200000000	0000002000		
		30GCCTTCGG	*00000000	10010010000	coccombar	CGGCCGGATG	
	65661 55661		50000000000000000000000000000000000000	MICHOCOMOD	00000001001	SMCCGCGTGG	
3.0	00/21	GGACGTTCCC	GCAGBCGGC6	ruubianiah	CCGCG:1cG		
30	66781	TOGACOTOGO	CGGGCTGCGC	CCCGGCGAGA	AGGICCITGAL	CCACGOGGCG	
	56841	TOGGOGGGGG	GGCCGTCCAG	ATCGCGCGGC	ATCTGGGCGC	GGAGGTGTAC	
	66901	GCGCCGCGAA	GCGCCATCTG	GTGGACCTGG	ACGGAGCGCA	TCTGGCCGAT	
	66961	CCGCGTTCGC	CGACGCGTTC	CCGCCGGTCG	ATGTCGTGCT	CHACTOGOTO	ACCGGTGAAT
	67021	TOOTOGACGO	greegregge	CTGCTCGCGG	CGGGTGGCCG	GTTCATCGAG	ATGGGGAAGA
35	27081	CGGACATCCG	CARCGARGEC	CLGCLGCCCC	TOGAROTGAT	GGACGCCGGC	COCGACCGGA
23			CAMCCMCCAC	cmccmccacc	TOTTOSCOCO	CGACGTGCTG	
	67141	. GOMOCOGA.	CMICGICGMG	CTGCTCGGCC	0001000000	CGGCTGGATG	
		CGGTCCACGC	CTGGGAGGTG	CGGCMGGCGC	occessor.		
	67261	GTCACACCGG	CAAGCTGGTG	CTGACGGTCC		SGATOCOGAG	
	67321	TCATCACCGG	CGGCTCCGGC	ACCCTCGCCG	GCATOCTOGO	DUGGCACCTG	
40	67381	ACACCTACCT	GCTCTCCCGC	ACCCCACCCC	CCGACACCAC	CCCCGGCACC	CACCTCCCCT
	67441	GCGACGTCGG	CGACCCCCAC	CAACTCGCCA	CCACCCTCGC	CCGCATCCCC	CAACCCCTCA
	57501	CCGCCGTCTT	CCACACGGGC	GGAACCCTCG	ACGACGCCCT	GCTCGACAAC	CTCACCCCCG
	27521	ACCGCGTCGA	CFCCCGGCGGC	AAACCCAAGG	CCGACGCCGC	CTGGCACCTG	CACCGGCTCA
	57.501	CCCCCGACAC	22700010010	CCCTTCCTCC	MOMENTON	agmagaagge	CTCATGGGCA
15	0.00	GCCCGGGGCA	0000110000	0001100100	ACCOCTTO	CALCOCCECTO	GCCGLACACC
45	0.00.		JA JAMOJUKE	GICGCGGCGA	modes modes	COMCOCCCC C	COCCUTATION
	57741	GCCGTGCGCA	AGGGCTGCCC	GCGCAG.CCC	ilulaidada	Chigidadea	GACGICAGCG
	67801	CGCTCACCGC	GARACTCACC	GACGLUGACC	COAGUGUAT	DUGGUGUAGU	GBAT LUCUGO
	57861	CGTTGAGCGC	CGCGGACGGC	ATGCGGCTGT	TOGACGCGGG	SACGCGTACC	CCGGAACCGG
	67921	TOGTOGTOGO	GACGACCGTC	GACCTCACCC	AGCTCGACGG	CGCCGTCGCG	COGTTGCTCC
50	67921	SOGGTOTOGO	CGCGCACCGG	GCCGGGCCGG	CGCGCACGGT	CBCCCGCAAC	GCCGGCGAAG
5 ()	85041	AGCCCCTGGC	CGTGCGTCTT	GCCGGGGGGTA	CCGCCGCCGA	GCAGCGGCGC	ATCATGCAGG
	501011	AGGTCGTGCT	00.0001011	CCCCCCCCCCCCC	madaamaaaa	acmagacaac	CGCGTGGCGG
	00101	CGGACCGTCC	COGCCACGCG	GCCGCGGGTCC	- TTCCCTC* C	CACCATCANO	CTGCGCAATC
	62101	CONTOCHE	GIICCGCGMG	CICGGIIICG	ALLOGUIGAU	Cacadata	CACCCATOCC
	69221	GGCTCGCGGC	CGAGACGGGG	CTGCGGCTGC	CUALUACUT	0010110AGC	CACCOCACCC
55	66261	CGGAGGCGCT	CACCGCCCAC	CTGCTCGACC	TGATUGACGC	LUCCACCGCC	UbaniuaCCG
	68341	GGGAGTCCCT	GCCCGCGGTG	ACGGCCGCTC	CCGTGGCGGC	CGCGCGGGAC	CAGGACGAGC
	62401	CERTCECCAT	CGTGGCGATG	GCGTGCCGGC	mgcccggmgg	TGTGACGTCG	CCCGAGGACC
	68.487 78.487	TGTGGCGGCT	CGTCGAGTCC	GGCACCGACG	CGATCACCAC	GCCTCCTGAC	GACCGCGGCT
	20501	GGGACGTCGA	00100.0100	GACGCGGLCC	CGGACGCGG	CGGCAAGGCG	TACAACCTGC
60	000771	GGGGCGGTTA	COUCCECCO	CACCACCACA	TOGRAGECCO	CTTCTTTCTC	ATORGTORGO
60	55581	abaddaGTTA			100000000	orrorrow.	5510000

	55641	203/2433337	COGCATGGAC	000000000000	::::::::::::::::::::::::::::::::::::::	CSAFACGGCG	TBGGAGGBBA
		7098020003	COSCATCACT	10000001000		GGAGGTCGGC	GTGTATGTGG
	33-3-	3130630060	GCAGGGGTAC	-336CT30333	COORDORACEC	CBASGGCCAC	GCGATCACCG
							3A33370000
5	49 8 8 C	00000000000000000000000000000000000000		- LUUUUNUUU. - maamaanaan	200000000000000000000000000000000000000	- GCTGCATCTG	
~	5594I	00000000000		GAMOTOSOTO	- 0100123.000 - 10000033838	- SGTCTCCGTA	GRETERE
				m.c.a.a.a.a.a.a.a			CTGAGTTCGC
	69001	0000000000	COTGGAGTTC	TO LUGULAGO	00000000000	33003A0GGG	CGCTGCAAGT
	69061	3377333333		addenicanca.		CGTGGGGGTG	DICGIACTEG
17.	99111		CORCOCCAC		why manager and	USUUSTOGTO	COCCCCACCC
10	R9181	CCGTCACGTC	03A0930300	TOCAACGGCO	12400000000	GAACGGGGCTC	TOGCAGCAGC
	69141	3337027000	GFA833033TC	5006056056	330T3A0033	nggggadgtg	GRIGTCGTCG
	69301	AGGGGGACGG	CACCGGCACC	0330103303	ACCCGGTCGA	93939A0 30 9	CTGCTCGCGA
	69361	COTACGOCA	33A203T003	GCACCGGTCT	3307333070	20T3MAGT03	AACATCGGAC
	A9421	ATGCCACSGC	0909330937	GTCGCGGGGG	TORTORAGAT	33 131A331 3	ATOGGOGGGG
15	2.0	CORTAGORICS	amennantu	Jr. 1370 J. 138	ASCCOTOBER	CGCCGTCGAC	TGGAGCACCG
	89841	SACASSTOTS	COTGCTCGGC	TOCAACCGGC	CCTGGCCGGA	COACGAGOGT	cogogogogo
	69661	0333037070	dedermodes	CTCAGCGGGA	CGAACGCGCA	CGTCATCCTG	GAACAGCACC
	69661	GTCCSGCGCC	CSTGGCGTCC	CAGCCGCCCC	ggccgccccg	TGAGGAGTCC	CAGCCGCTGC
	69721	CGTGGGTGCT	crecacacaca	ACTCCGGCCG	cgcracagac	COAGGCGGCC	CGGCTGCGCG
20	69781	ACCACCTOSC	GGCGGCACCG	GACGCGGATC	CGTTGGACAT	CGGGTACGCG	CTGGCCACCA
~~	69841	GCCGCGCCCA		CGTGCCGCGG	TOGTOGOGAC	CACCCCGGAC	GGATTCCGTG
	69901	COSCSCTOSA	CGGGGTGGGG	GACGGCGCGG	AGGCGCCCGG	AGTOGTOACC	GGGACCGCTC
			CGTCGCCTTC	CTCTTCGACG	300A000000	004606060	GGAATGGGGC
	69961	AGGAGCGGCG		CCCGTCTTCG	000000000000000000000000000000000000000	GGACGAGGTC	TOOGACGOGT
25	70021	GCGAGCTCCA	COGCOGGTTC				
£3	70081	TOGGCAAGCA	CCTCAAGCAC	TCCCCCACGG	ACGTCTACCA	CGGCGAACAC	GGCGCTCTCG
		CCCATGACAC	CCTGTACGCC	CAGGCCGGCC	TGTTCACGCT	CGAAGTGGCG	CTGCTGCGGC
	70201	TGCTGGAGCA		CGGCCGGACG	TGCTCGTCGG	GCACTCCGTC	GGCGAGGTGA
	70261	COGCGGCGTA	CGCGGCGGGG	GTGCTCACCC	TGGCGGACGC		ATCGTGGCCC
	70321	GGGGGGGGGC	GCTGCGGGGG		GGGCGATGCT	CGCCGTCGAC	
30	70381	CGGAGGTCGG			TOGCCGCGGT	CAACGGCCCG	TOCGCCGTGG
	70441	TGCTCGCCGG	TTCGCCGGAC		CGTTCGAACG	GGAGTGGTCG	
	70501	GGCGCACGAA			CGTTCCACTC		GACGGTGCGC
	70561	TCGACGGCTT				cacaaaaacaa	CTGCCGGTGG
	70621	TGTCCACGAC	GACGGGGCGG	GACGCCGCGG	ACGACCTCAT	AACGCCCGCG	CACTGGCTGC
35	70681	GCCATGCGCS	TOGGCCGGTG	CTGTTCTCGG	ATGCCGTCCG	GGAGCTGGCC	GACCGCGGCG
	70741	TCACCACGTT	CGTGGCCGTC	GGCCCCTCCG	GCTCCCTGGC	GTCGGCCGCG	GCGGAGAGCG
	70801	COGGGGAGGA	CGCCGGGACC	TACCACGCGG	TGCTGCGCGC	CCGGACCGGT	GAGGAGACCG
,	70861	CGGCGCTGAC	CGCCCTCGCC	GAGCTGCACG	CCCACGGGGT	CCCGGTCGAC	CTGGCCGCGG
	70921	TACTGGCCGG	TGGCCGGCCA	GTGGACCTTC	CCGTGTACGC	GTTCCAGCAC	CGTTCCTACT
40	70981	GGCTGGCCCC	GGCCGTGGCG		CCACCGTGGC	GGACACCGGG	GSTCCGGCGG
10				ACCGTCGCCG		TOGGOGGAGO	GCGGCGCTGC
	71101	TOGGOGTOAC		GACGTCGATG		GTTCTTCGCG	
	71101	TOBOOTOMS	CCTCCCCCCCC	CTCCCCANCC	a contractor	GGCAACCGGG	CTGGACCTGC
	1101	ACTUMU.GGC	CCTCTTCCCCC	CIGCOCIACC	necarecer.	CACCGCGTTC	CTCCAGGLCC
4.5	71221		0.101.101.00	CACCACCACA	COGCCGCGCI	CGACGACGCG	CICCAGGMCC
45	11281	GDATCGAGGC	CGGCCAGGAC	CGGAICGAGG	* CCGGCGAGGA	CATOCACCC	T COROCCOCC
	71341	TCTCGCT TCT	GGAGGAGALG	GAGICGUICG	MODELECT MANAGEMENT	CATCGCGGCG	ALGCEGGECC
	71401	OGGAGCGTGC	GGCCATCGCC	GATCTGCTCG	ACAAGCICGC	CCATACCIGG	MAGGACIACE
	71461	GATGAGCACC	GATACGCACG	AGGGAACGCC	00000000000	CGCTGCCCAT	TCGCGATCUA
	71521	GGACGGTCAC	CGCGCCATCC	TGGAGAGCGG	CACGGTTGGGT	TCGTTCGACC	TGTTCGGCGT
50	71581	CAAGCACTGG	CIGGICGCCG	CCGCCGAGGA	CGTCAAGCTG	GTCACCAACG	ATCCGCGGTT
	71641	TAGETEGGEE	GCGCCGTCCG	AGATGCTGCC	CGACCGGGGG	CCCGGGCTGGT	TCTCCGGGAT
	71701	GGACTCACCG	GAGCACAACC	GCTACCGGCA	GAAGATCSCG	GGGGACTTCA	CACTGCGCGC
	71761	GGCGCGCAAG	CGGGAGGACT	TOGTOGCOGA	GGCCGCCGAC	GCCTGCCTGG	ACGACATEGA
	71821	GGCCGCGGGA	CCCGGCACCG	ACCTCATCCC	CGGGTACGCC	AAGCGGCTGC	CCTCCCTCGT
55	71881	CATCAACGCG	CTGTACGGGC	TCACCCCTGA	GGAGGGGGCC	GTGCTGGAGG	CACGGATGCG
_	7]941	CGACATCACC	GGCTCGGCCG	ATCTGGACAG	CGTCAAGACG	CTGACCGACG	ACTTCTTCGG
	25001	0030000000	CGGCTGGTCC	GCGCGAAGCG	TGACGAGCGG	GGCGAGGACC	TGCTGCACCG
	72061	GOTGGCCTOC	6068048008	GCGAGATCTC	GCTCAGCGAC	GACGAGGCGA	CGGGCGTGTT
	72121	- CGCGACGCTCG	CTGTTTCGTCG	GCCACGACTC	GGTGCAGCAG	ATGGTCGGCT	ACTGCCTCTA
60	70101	CSCSACTCOMC	VOCUTUCUUS	AGCAGCAGGC	GGCGCTGCGC	GCGCGCCCGG	AGCTGGTCGA
Ç.	/4101	190M010010	1,000,10000				

		044000000000	GAGGAGATGC	2000 C	0000000000	CAGATGGGGG	TACCCC3037
						GIGGGGGACA	
			GMCG.JJMIC				
	71361	GOTOTACTOS	ACGGCCAACU	GCGACCCGA		CABCCCGACA	CCTTCGATGT
	*** : * : * :	GACGTGCCG	000000000000000000000000000000000000000	ACTICGCOCCT:	0.330.50	ATTORORAGE	370003355A
5				TORREGIOSS	~~ ~~ ~~ ~~ ~~		
Σ,			CGGGTGGTGA			uanac	STOTECCCGGA
	70541	0000033373	GCCGGCGACG	TOCCGATGAA	a salatik ada salad ada salad sa	333 37777 77	30000330 03 A
	70601		ACCTGGGGGG	CGCCATGAGT	a <u>r</u> agagamas	AGACGTTGC3	STTSSSSAAC
						TOOTOTAGEG	
	ี โซซิโ	GGGACGACGG		CAACGCGGGG	35.30 - 345.5	IAUUU	GGAGATOTTC
		ACCORGOSCT	GCTACCTGCG	CCACGGTGTT	3/10/2000	2000 0A 00T	SSTGTTCGAC
10	7,781	GTCGGGGGGA		GTTCACGCTT		733A3T3T00	TEGTGTGACC
1 %							-
	72841	GTGCACGCCT	m	GCCCGTGCCG	1.54644	7306630 37 A ACAGCTCCGG	CGTGACGCGG
	70901	CACGGCATIC	CGGGCCAGGC	GGACCAGTGI	-3033TOT 323	ACAGCTCCGG	CROCCGGRAG
	72981	ATGACOTTCT		CACGCTGATG	عديمهم ومحديد	ACCCGARGE	codesectes
					;5533		
	73021	ACGBAGCTCT	TGCGCACGCT	CGGCCTCAAC	630330TATA	COGOCGAGGA	CGTCGACACC
15	77081	ATGCTCGCGC	AACTGCCCGA	CGTCASCGAG	GAGATOGAWA	000000000000	0066670700
	73141	GACGICATCS	CGGAGCGCGG	mampon on an	Are a grandming in	134A33 7 03A	CGTGGAGAAG
				.MICOMOSSO	71.000000101	100000000	
	73201	AGCGAACGGC			GACACCOACT		CCGCCAGGTC
	73261	9709099AGG	mananaanam	chacgedges	CTCGAGGAGA	TOOMOROGOT	GCTCCGCGGC
	73321	CATOSOTTOR				COGGCACGGG	CATCCACCAG
200				CGAGCAGGAA		1003CA56GG	
20	73381	9709009090	GGCGGGTT3GC	CGGCTGAGCG	0097039930	00000000000000000000000000000000000000	CGCACCGGCG
	73441	GCCGCGGTGC	GGACGGCGGC	TCAGCCGGCG	TOBGACAGTT	COCOGGGGGGAG	TTGCTGACGG
					STGAGGTBOT	257221000	GOTGGAGGTG
	73501				olomabla		
	73861	ACGAACAGCT	GGCTGGCGAT	CTCCTTGTTG	GTGCGCCCCA	CCGCGGGGTG	CGACGCCACC
	73621			CGATGTGATC	0307303000	30000250000	CTG33TGCCG
25							
25	13987	TCCGCGTCCG	AGGACTCCCC	AUCUMBUUCH.	CGGAGGAGCS	achebbe the	GCACTGGGTC
	73741	GCGAGGTGCC	STGCGCGGCG	GAACAGTCCC	CGCCCACGGC	ToTGCCGCCG	GAGCATGCCG
	73601	CACGCTTCGC	CONTRACTOR	abagbeacas.	GOOD GODOGE	20000000000	GCACATGATG
			CCM101C33C	07007.00000	> maccoammac	0000000000	
	73561			GAGCAGTTCG		CCGGCGGACT	STAGGCCGCC
	73921	TOCACCCCCA	GOGTCATCAC	CCGCGCCCCGG	GACCCCATCG	GCCGGGGACAG	CTGCTCGGAG
30		ATGAGCCTCA				35.53030000	GGCGGCGTCG
20	3301	michael.cm	0100010010	ACCCCCCCCCC	. 2021 201 20	CD 2021 MCCC	
	74041	ACCCGCCACA	GGGCCAGGCC	CGGCACGTCG	Milbadobas	GTOGGATCCG	
	74:01	TCCCGGAACG	CGTTGTACGC	CGCCCGGTAC	CGCCCGGCCG	CGAGATGGTG	TTGCCCACGG
		GCCCAGACCA			COGGRESTOR	COTOCGGCAA	CARCTEGGER
		Grandwarder	id. GCMGibb	GAAGAGGC TO			
	74221	AGCCACCGCT	CCGCCCGGTC	CAGGTCGCCC	Ralldaami da	CGGCGGCCAC	
35	71791	AGCGGCAATG	CGGCGGCCAT	CCCCCASGAS	SBOACGACCE	3333333C3AG	CGCGGCCTCG
	74513	CCGCATTCGA	CCCCCCCCCC			19301 1036 0	
	/9391	CCGUATICGA		CHORICOCCA	0000000000		
	74401	GCCTGGACCG	CCTCGTCGGC	CGGGGTCCGC	ATGTTGTCGT	CACCGGGCCAG	CTTGTCGACC
	73381	CASSACTSSA	COCCETORS	GTCCTCGGCG	TAGAGCAGGG	CONSCAROSC	CATCATGGTC
					TGGAGCACGT		GGCCTCGGGC
	74521	GIGGICCGGI	JUSTUSTOAU	CCGGGAGTGC			
40	74581	TGTTCGGACC	AGCCGCGCAG	CGCGTTGCTC	AGGGCCTTGT	CGGGGGACGGC	
	71611	ACGGCTCCGG	MILLOGRAGAC	GACCTCGTCC	maggaaggaa	GATCGGCCGG	ACGCGGCGGA
	74341	TOGGCCGCGC	TARRESTANT.	02.000000000	are reameer.	2010000000	cmeedeecas
	74-01	TCGGUUGGCGC	CGGGATAGAT	CHECECONIC	GMUMUGICUU	COMCOUGONS	2190000000
	74761	CCCTGCTCGC	TOGGGGGGGGC	GGAGCGCTGG	GCCGCCAGGA	CCTCGGCGGC	CTCGCCCGGC
	77601	CGCCCGTCCA	modeceageda	GCAGGCGAGC	GACACGGGGT	GCTCGCTGGA	GAGGAGCCGT
4.5	14021		TOOCCAOCCA	20.100001100	meacecca	*******	rmedetalac
45	74851	TCCCSUGACG	CGGTGAGCAG	CICEGGGACA	100000000	M.01000000	ATOGCAGAGC
	74941	CGCTCGATGG	CGGCGGTGTC	GACGCGCAGT	GCGGCGTGGA	CGGCGGGGTC	GTCGGAGGCC
	75001	CGGTA.BUCGA	a Caraca Coda	COTCACCCC	modmodalaam	ngoogogoag	GTGGTGCTCG
	- 3002	00011.0303M	Mercenoorn	000000000000	magaaaaaa	0010000000	CCE FOCUL INC
	75061	CGCGCGGCGT	CGGTGAACAG	CCCGGGCGACG	100000000	JUNCEUBBUC	GGIMCCCMIC
	فرنعك	TGGTGGCGGG	CSASCACCTT	GCTGGCCACG	CCCCGGGTCCC	GURGURGTTO	CAGCGCCAGC
50	2222	TOGTGCAGGC	OT COCCEPT	coccacacas	10000000000	GTACGACGGA	GCGGGGCCGCG
10	2.51	. Caralness	CWCGCCGCTC	00000000000	neoneoneon.	07:00:00:00:00	omenenenen
	75041	GGGTGCGGGA	ACCGCCCTTC	CCGCAGCAGC	Locutor Lan	oundare 110	0.00000.00
	74501	TOGACCGCCT	CGGTGTCGAG	GCCGGTCATC	CGCTGGACGA	GGGTGAGTTC	GACACTCTCS
		CCGAGCACGG	00010100110	acceptoceme.	tiangnaanna	ggaagassaa	202022200220
	12327	Chambentag	COURAGE TELE	GOCOMEON - C	7/00/30/30/00/	Journal Commod	moreomecon
	75421	CCGAGGTAGG	CGAGCCGGTA	CGCCGGCCCC	GCGACCACTT	LUAGGCACCC	i ombolfbool
55	75 121	GTCCGTGCCT	COCCCATOR	GTCGATCAGG	COGTGGCCCA	GGAGCAGGTT	GCCGCCGGTC
~' . '	7401	0.000.000.	0.0000011010	0100.1101.00	wcccccmcam	0000000000	ccadactea
	75541	GCCCGGAACG	CCTGGGCCAC	CACGTCGTCG	100000.001	01000010	
	75601	AGTTCGGTGG	TOTGOGGGTG	GGTGAGCGGG	CGCAGCGCGA	TOTOOTGGTA	GTGGCGCAGA
	757	CTCAGCAGTG	0000000000	TTGGGLGTGG	GOGGGGGGTCG	GCCGGAGCAG	CTCGGTCAGC
	10001	しょしかい かっし	UL GC CCGMM	1100000000	0000000000	2012012001	accerterer.
	75721	ACGATGGCGA	CACGGGCCCG	GUTGATGUGG	Techecology	MJERJUMGU A	GUGURGUAG
60	75767	GGCGCGTCGG	CGTGGTGCAC	GTCGTCGATG	CCGATCAGTA	CGGGCCGCTC	CGCGGCGAGC
Ç/ C-	J - L =						

40

45

46

TREAS STORESERY) TECEMORERS TOOSTOOCS ACCESTIST DEACGTOSEC CEGCAGGITT NEAR TOSSAGES STORAGES SASCASCTOT GENEROSS DESCRIPTO GENEROSS TERRI AGGAGGTGG DGAGGATGC GTAGGGGAGG GCCGGTTCT DGATGGAGGA GACGGGGGA VALL AGGGTGAGA AGCGGGGTT GGCGGCGGG GCGTTGGAGGA 74TGGGAGGA GACGGGGGA TAGGG ATGGGCCGG TGAGGGGGG GAGGAGGCG GGCGGGCGC GCGTTGGAGGGT GAGGGCGGG TRIBLE TERMESGAAD OBAACTOSTO ATOSCESSES ATOAGSTOTE SUGGAGATAA GOGOGOTATO TRIBLE ACGAATGGAA OTACCTOSCS ACCETESTEG AAACCOATAR SCATCACATS SETTETTGAT "ASAS CRACCAST SESSACCOS TOCTOSTOS LESCALA AGATGORAS SOLIGITAS "ASAS CRACCASTOS CONTRACTORAS CASCAGAS COCOCASCAS AGATGACATA AGACCATA AGACCATA AGACCATA AGACCATA AGACCATA AGACCAGAGAA AAAATGACCAG CAGGAGAAGAA 10 neggi loadonatii garandosa lottaatus tunoanotta minaratosa tusasanan neggi pamanatist garandosa graffatas assacanas tuatagaat uroangas assacanas nessi mattagana graffatas garandosa tusatagaan uroangas assacanas attagaaga nessi abbandasta garandosaa garantast assacana assacanas attagaagas TAGEL GOAGOTOSTO STOLLAGGOS AGGTGGTOOT GALAGGOS COGOTOSTO GOGOTOGGA GOGOGOGO 15 REEL GUAGOTUSTO STOLLAUGUS AGRIGOTUT GELUGOSUS LIGUTEUGAU GUUGOOGOO MERAI GEOOLOATEOT STOOLOGAAC GOCACOGGGA GUUGOTUGA UURUUGAAC GUGAGTOTO MERAI DECOCATEOT STOOLOGAAC AGCACOAGGG GACGGTOCAS UURUUGGGTO AACGCCTOGG VERAI DECOGGGGTA CTGCACGGGG TACACGTOGG CCACGGGGG GAGGGCACGG GCCAGGGGAA 20 TERRI GGTAGAARR DESCOGATICO COGGGGGTSS SCASCASCAS DARROGGTACO GGGGGCGAG
VUNAL GOGTGGGGAA GAACTGCGG AGCCAGAGTT COGAGGTCAD DRIROCGTACO GGGCGGGGC
TILL CTGGGGAGCS CGGAACCGGG TGATCTCGSC CAAGTGCTT TOTCGGATCT CCGGGTCGGT This CACGOCCAT CONTOCTOOG GOGCCAGACA GAGGACGOOG ACTITICOGT TOTGCACATT 77221 GOGATGACA TOGOGCACOG COGACOCGAC GAGGACGOO GAGGACGOO GOGACAGCGT
TO221 GOGATGACA TOGOGCACOG COGACOCGAC GTOGOCTIC CACGCCTAC GATAGTTCGC
TO281 CGGGTGCACC ATCCCCTTGC AGATCACGCG GTTCGCCTIC CACGCATCAC GATAGTTCGC
77341 GAAGTGGGTA CCGATGATCC GCTTCACGGA CATCCACAGG TACCGATTGT CAAAGGCGTG
77401 CTCGTATCCC GAGGTTGACG CGCAGGTGAC GATCGTGCCA CCCGAGGTG TCACGTAGAC 25 27461 ACTOGOGOG ARCSTOGOGO GCCCCGGGTG CTCGARCAGS ATSTOGGGGAT CGTCACCGGC TIS21 GGTCAGCTCC CGGATC 30

Those of skill in the art will recognize that, due to the degenerate nature of the genetic code, a variety of DNA compounds differing in their nucleotide sequences can be used to encode a given amino acid sequence of the invention. The native DNA sequence encoding the FK-520 PKS of *Streptomyces hygroscopicus* is shown herein merely to illustrate a preferred embodiment of the invention, and the present invention includes DNA compounds of any sequence that encode the amino acid sequences of the polypeptides and proteins of the invention. In similar fashion, a polype side can typically tolerate one or more amino acid substitutions, deletions, and insertions in its amino acid sequence without loss or significant loss of a desired activity. The present invention includes such polypeptides with alternate amino acid sequences, and the amino acid sequences shown merely illustrate preferred embodiments of the invention.

The recombinant nucleic acids, proteins, and peptides of the invention are many and diverse. To facilitate an understanding of the invention and the diverse compounds and methods provided thereby, the following general description of the FK-520 PKS genes and modules of the PKS proteins encoded thereby is provided. This general description is followed by a more detailed description of the various domains and modules of the FK-520

10

15

20

25

30

PCT/US99/22886

PKS contained in and encoded by the compounds of the invention. In this description, reference to a heterologous PKS refers to any PKS other than the FK-520 PKS. Unless otherwise indicated, reference to a PKS includes reference to a portion of a PKS. Moreover, reference to a domain, module, or PKS includes reference to the nucleic acids encoding the same and vice-versa, because the methods and reagents of the invention provide or enable one to prepare proteins and the nucleic acids that encode them.

The FK-520 PKS is composed of three proteins encoded by three genes designated fkb.4. fkbB, and fkbC. The fkb.4 ORF encodes extender modules 7 - 10 of the PKS. The fkbB ORF encodes the loading module (the CoA ligase) and extender modules 1 - 4 of the PKS. The fkbC ORF encodes extender modules 5 - 6 of the PKS. The fkbP ORF encodes the NRPS that attaches the pipecolic acid and cyclizes the FK-520 polyketide.

The loading module of the FK-520 PKS includes a CoA ligase, an ER domain, and an ACP domain. The starter building block or unit for FK-520 is believed to be a dihydroxycyclohexene carboxylic acid, which is derived from shikimate. The recombinant DNA compounds of the invention that encode the loading module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of methods and in a variety of compounds. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 loading module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for the loading module of the heterologous PKS is replaced by the coding sequence for the FK-520 loading module, provides a novel PKS coding sequence. Examples of heterologous PKS coding sequences include the rapamycin. FK-506, rifamycin, and avermectin PKS coding sequences. In another embodiment, a DNA compound comprising a sequence that encodes the FK-520 loading module is inserted into a DNA compound that comprises the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the loading module coding sequence is utilized in conjunction with a heterologous coding sequence. In this embodiment, the invention provides, for example, either replacing the CoA ligase with a different CoA ligase, deleting the ER, or replacing the ER with a different ER. In addition, or alternatively, the ACP can be replaced by another ACP. In similar fashion, the corresponding domains in another loading or extender module can be replaced by one or more domains of the FK-520 PKS. The resulting heterologous loading module coding sequence can be utilized in conjunction

with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide.

5

 $\mathbf{I}()$

15

20

25

30

The first extender module of the FK-520 PKS includes a KS domain, an AT domain specific for methylmalonyl CoA, a DH domain, a KR domain, and an ACP domain. The recombinant DNA compounds of the invention that encode the first extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 first extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the first extender module of the FK-520 PKS or the latter is merely added to coding sequences for modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the first extender module of the FK-520 PKS is inserted into a DNA compound that comprises the remainder of the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, all or only a portion of the first extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2hydroxymalonyl CoA specific AT; deleting either the DH or KR or both; replacing the DH or KR or both with another DH or KR; and/or inserting an ER. In replacing or inserting KR, DH, and ER domains, it is often beneficial to replace the existing KR, DH, and ER domains with the complete set of domains desired from another module. Thus, if one desires to insert an ER domain, one may simply replace the existing KR and DH domains with a KR, DH, and ER set of domains from a module containing such domains. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a gene for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous first extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous

10

15

20

25

30

WO 00/20601 PCT/US99/22886

PKS can be replaced by one or more domains of the first extender module of the FK-520 PKS.

In an iliustrative embodiment of this aspect of the invention, the invention provides recombinant PKSs and recombinant DNA compounds and vectors that encode such PKSs in which the KS domain of the first extender module has been inactivated. Such constructs are especially useful when placed in translational reading frame with the remaining modules and domains of an FK-520 or FK-520 derivative PKS. The utility of these constructs is that host cells expressing, or cell free extracts containing, the PKS encoded thereby can be fed or supplied with N-acylevsteami — thioesters of novel precursor molecules to prepare FK-520 derivatives. See U.S. patent application Serial No. 60/117,384, filed 27 Jan. 1999, and PCT patent publication Nos. US97/02358 and US99/03986, each of which is incorporated herein by reference.

The second extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the second extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 second extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the second extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the second extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, all or a portion of the second extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a maionyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting the KR and/or the inactive DH; replacing the KR with another KR; and/or inserting an active DH or an active DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these

10

15

20

25

30

replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous second extender module coding sequence can be utilized in conjunction with a coding sequence from a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the second extender module of the FK-520 PKS.

The third extender module of the FK-520 PKS includes a KS, an AT specific for malonyl CoA, a KR, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the third extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 third extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the third extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the third extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, all or a portion of the third extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the malonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting the KA and/or the inactive DH; replacing the KR with another KR; and/or inserting an active DH or an active DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous third extender module coding sequence can be utilized in conjunction with a coding sequence from a PKS that synthesizes FK-520, an

25

30

FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the third extender module of the FK-520 PKS.

PCT/US99/22886

The fourth extender module of the FK-520 PKS includes a KS, an AT that binds 5 ethylmalonyl CoA, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the fourth extender module of the FK-520 PKS and the corresponding polypentides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 fourth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS 10 is either replaced by that for the fourth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the fourth extender module of the FK-520 PKS is inserted into a 15 DNA compound that comprises the remainder of the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the fourth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the ethylmalonyl CoA specific AT with a malonyl CoA, methylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; and/or deleting the inactive DH, inserting a KR, a KR and an active DH, or a KR, an active DH, and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, a PKS for a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous fourth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the fourth extender module of the FK-520 PKS.

As illustrative examples, the present invention provides recombinant genes, vectors, and host cells that result from the conversion of the FK-506 PKS to an FK-520 PKS and vice-versa. In one embodiment, the invention provides a recombinant set of FK-506 PKS

15

20

25

30

PCT/US99/22886

genes but in which the coding sequences for the fourth extender module or at least those for the AT domain in the fourth extender module have been replaced by those for the AT domain of the fourth extender module of the FK-520 PKS. This recombinant PKS can be used to produce FK-520 in recombinant host cells. In another embodiment, the invention provides a recombinant set of FK-520 PKS genes but in which the coding sequences for the fourth extender module or at least those for the AT domain in the fourth extender module have been replaced by those for the AT domain of the fourth extender module of the FK-506 PKS. This recombinant PKS can be used to produce FK-506 in recombinant host cells.

Other examples of hybrid PKS enzyman f the invention include those in which the AT domain of module 4 has been replaced with a malonyl specific AT domain to provide a PKS that produces 21-desethyl-FK520 or with a methylmalonyl specific AT domain to provide a PKS that produces 21-desethyl-21-methyl-FK520. Another hybrid PKS of the invention is prepared by replacing the AT and inactive KR domain of FK-520 extender module 4 with a methylmalonyl specific AT and an active KR domain, such as, for example, from module 2 of the DEBS or oleandolide PKS enzymes, to produce 21-desethyl-21-methyl-22-desoxo-22-hydroxy-FK520. The compounds produced by these hybrid PKS enzymes are neurotrophins.

The fifth extender module of the FK-520 PKS includes a KS, an AT that binds methylmalonyl CoA, a DH, a KR, and an ACP. The recombinant DNA compounds of the invention that encode the fifth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 fifth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the fifth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS. In another embodiment, a DNA compound comprising a sequence that encodes the fifth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the fifth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA

15

20

25

30

specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting any one or both of the DH and KR; replacing any one or both of the DH and KR with either a KR and/or DH; and/or inserting an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous fifth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes rK = an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the fifth extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the DH domain of the fifth extender module have been deleted or mutated to render the DH non-junctional. In one such mutated gene, the KR and DH coding sequences are replaced with those encoding only a KR domain from another PKS gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-19 to C-20 double bond of FK-520 and has a C-20 hydroxyl group. Such analogs are preferred neurotrophins, because they have little or no immunosuppressant activity. This recombinant fifth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this fifth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (lacking the C-19 to C-20 double bond of FK-506 and having a C-20 hydroxyl group) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the DH domain of module 5 has been deleted or otherwise rendered inactive and thus produces this novel polyketide.

The sixth extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds

WO 00/20601 PCT/US99/22886

of the invention that encode the sixth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 sixth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the sixth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the sixth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

10

15

20

25

30

In another embodiment, a portion of the sixth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting any one, two, or all three of the KR, DH, and ER; and/or replacing any one, two, or all three of the KR, DH, and ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous sixth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the sixth extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the DH and ER domains of the sixth extender module have been deleted or mutated to render them non-functional. In one such mutated gene, the KR, ER, and DH coding sequences are replaced with those encoding only a KR domain from another PKS gene. This can also be accomplished by simply replacing the coding sequences for extender module six with those for an extender module having a methylmalonyl specific AT and only a KR domain from a heterologous PKS gene, such as,

10

15

20

25

30

for example, the coding sequences for extender module two encoded by the *ervAI* gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that has a C-18 hydroxyl group. Such analogs are preferred neurotrophins, because they have little or no immunosuppressant activity. This recombinant sixth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this sixth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (having a C-18 hydroxyl group) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the DH and ER domains of module 6 have been deleted or otherwise rendered inactive and thus produces this novel polyketide.

The seventh extender module of the FK-520 PKS includes a KS, an AT specific for 2-hydroxymalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds of the invention that encode the seventh extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 seventh extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the seventh extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encod the seventh extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion or all of the seventh extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the 2-hydroxymalonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or

15

20

25

30

3,

malonyl CoA specific AT: deleting the KR, the DH, and or the ER; and/or replacing the KR, DH, and/or ER. In addition, the KS and or ACP can be replaced with another KS and or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous seventh extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the seventh extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the AT domain of the seventh extender module has been replaced with those encoding an AT domain for malonyl, methylmalonyl, or ethylmalonyl CoA from another PKS gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-15 methoxy group, having instead a hydrogen, methyl, or ethyl group at that position, respectively. Such analogs are preferred, because they are more slowly metabolized than FK-520. This recombinant seventh extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this seventh extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (C-15-desmethoxy) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the AT domain of module 7 has been replaced and thus produces this novel polyketide.

In another illustrative embodiment, the present invention provides a hybrid PKS in which the AT and KR domains of module 7 of the FK-520 PKS are replaced by a methylmalonyl specific AT domain and an inactive KR domain, such as, for example, the AT and KR domains of extender module 6 of the rapamycin PKS. The resulting hybrid PKS produces 15-desmethoxy-15-methyl-16-oxo-FK-520, a neurotrophin compound.

15

20

25

30

The eighth extender module of the FK-520 PKS includes a KS, an AT specific for 2-hydroxymalonyl CoA, a KR, and an ACP. The recombinant DNA compounds of the invention that encode the eighth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 eighth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the eighth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the eighth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the eighth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the 2-hydroxymalonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or malonyl CoA specific AT; deleting or replacing the KR; and/or inserting a DH or a DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous eighth extender module coding sequence can be utilized in conjunction with a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the eighth extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the AT domain of the eighth extender module has been replaced with those encoding an AT domain for malonyl, methylmalonyl, or ethylmalonyl CoA from another PKS gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-13 methoxy group, having instead a hydrogen, methyl, or ethyl group at that position, respectively. Such

WO 00/20601 PCT/US99/22886

analogs are preferred, because they are more slowly metabolized than FK-520. This recombinant eighth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this eighth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (C-13-desmethoxy) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the AT domain of module 8 has been replaced and thus produces this novel polyketide.

5

10

15

20

25

30

The ninth extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds of the invention that encode the ninth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 ninth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the ninth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the ninth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the ninth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting any one, two, or all three of the KR, DH, and ER: and/or replacing any one, two, or all three of the KR, DH, and ER with another KR, DH, and/or ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can

10

15

20

25

30

originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous ninth extender module coding sequence can be utilized in conjunction with a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the ninth extender module of the FK-520 PKS.

The tenth extender module of the FK-520 PKS includes a KS, an AT specific for malony! Come if an ACP. The recombinant DNA compounds of the invention that encode the tenth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 tenth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the tenth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the tenth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion or all of the tenth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the malonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; and/or inserting a KR, a KR and DH, or a KR, DH, and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous tenth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a

10

15

20

25

30

60

module of a heterologous PKS can be replaced by one or more domains of the tenth extender module of the FK-520 PKS.

The FK-520 polyketide precursor produced by the action of the tenth extender module of the PKS is then attached to pipecolic acid and cyclized to form FK-520. The enzyme FkbP is the NRPS like enzyme that catalyzes these reactions. FkbP also includes a thioesterase activity that cleaves the nascent FK-520 polyketide from the NRPS. The present invention provides recombinant DNA compounds that encode the fkbP gene and so provides recombinant methods for expressing the fkbP gene product in recombinant host cells. The recombinant jkbP genes of the invention include those in which the coding sequence for the adenviation domain has been mutated or replaced with coding sequences from other NRPS like enzymes so that the resulting recombinant FkbP incorporates a moiety other than pipecolic acid. For the construction of host cells that do not naturally produce pipecolic acid, the present invention provides recombinant DNA compounds that express the enzymes that catalyze at least some of the biosynthesis of pipecolic acid (see Nielsen et al., 1991, Biochem. 30: 5789-96). The fkbL gene encodes a homolog of RapL, a lysine cyclodeaminase responsible in part for producing the pipecolate unit added to the end of the polyketide chain. The fkbB and fkbL recombinant genes of the invention can be used in heterologous hosts to produce compounds such as FK-520 or, in conjunction with other PKS or NRPS genes, to produce known or novel polyketides and non-ribosmal peptides.

The present invention also provides recombinant DNA compounds that encode the P450 oxidase and methyltransferase genes involved in the biosynthesis of FK-520. Figure 2 shows the various sites on the FK-520 polyketide core structure at which these enzymes act. By providing these genes in recombinant form, the present invention provides recombinant host cells that can produce FK-520. This is accomplished by introducing the recombinant PKS. P450 oxidase, and methyltransferase genes into a heterologous host cell. In a preferred embodiment, the heterologous host cell is *Strep.omyces coelicolor* CH999 or *Streptomyces lividans* K4-114, as described in U.S. Patent No. 5,830,750 and U.S. patent application Serial Nos. 08/828,898, filed 31 Mar. 1997, and 09/181,833, filed 28 Oct. 1998, each of which is incorporated herein by reference. In addition, by providing recombinant host cells that express only a subset of these genes, the present invention provides methods for making FK-520 precursor compounds not readily obtainable by other means.

In a related aspect, the present invention provides recombinant DNA compounds and vectors that are useful in generating, by homologous recombination, recombinant host

15

20

25

30

PCT/US99/22886

cells that produce FK-520 precursor compounds. In this aspect of the invention, a native host cell that produces FK-520 is transformed with a vector (such as an SCP2* derived vector for Streptomyces host cells) that encodes one or more disrupted genes (i.e., a hydroxylase, a methyltransferase, or both) or merely flanking regions from those genes.

When the vector integrates by homologous recombination, the native, functional gene is deleted or replaced by the non-functional recombinant gene, and the resulting host cell thus produces an FK-520 precursor. Such host cells can also be complemented by introduction of a modified form of the deleted or mutated non-functional gene to produce a novel compound.

In one important embodiment, the present invention provides a hybrid PKS and the corresponding recombinant DNA compounds that encode those hybrid PKS enzymes. For purposes of the present invention a hybrid PKS is a recombinant PKS that comprises all or part of one or more modules and thioesterase cyclase domain of a first PKS and all or part of one or more modules, loading module, and thioesterase cyclase domain of a second PKS. In one preferred embodiment, the first PKS is all or part of the FK-520 PKS, and the second PKS is only a portion or all of a non-FK-520 PKS.

One example of the preferred embodiment is an FK-520 PKS in which the AT domain of module 8, which specifies a hydroxymalonyl CoA and from which the C-13 methoxy group of FK-520 is derived, is replaced by an AT domain that specifies a malonyl, methylmalonyl, or ethylmalonyl CoA. Examples of such replacement AT domains include the AT domains from modules 3, 12, and 13 of the rapaymycin PKS and from modules 1 and 2 of the erythromycin PKS. Such replacements, conducted at the level of the gene for the PKS, are illustrated in the examples below. Another illustrative example of such a hybrid PKS includes an FK-520 PKS in which the natural loading module has been replaced with a loading module of another PKS. Another example of such a hybrid PKS is an FK-520 PKS in which the AT domain of module three is replaced with an AT domain that binds methylmalonvl CoA.

In another preferred embodiment, the first PKS is most but not all of a non-FK-520 PKS, and the second PKS is only a portion or all of the FK-520 PKS. An illustrative example of such a hybrid PKS includes an erythromycin PKS in which an AT specific for methylmalonyl CoA is replaced with an AT from the FK-520 PKS specific for malonyl CoA.

Those of skill in the art will recognize that all or part of either the first or second PKS in a hybrid PKS of the invention need not be isolated from a naturally occurring source. For example, only a small portion of an AT domain determines its specificity. See U.S. provisional patent application Serial No. 60.091,526, incorporated herein by reference.

The state of the art in DNA synthesis allows the artisan to construct *de novo* DNA compounds of size sufficient to construct a useful portion of a PKS module or domain. For purposes of the present invention, such synthetic DNA compounds are deemed to be a portion of a PKS.

Thus, the hybrid modules of the invention are incorporated into a PKS to provide a hybrid PKS of the invention. A hybrid PKS of the invention can result not only:

- (i) from fusions of heterologous domain (where heterologous means the domains in that module are from at least two different naturally occurring modules) coding sequences to produce a hybrid module coding sequence contained in a PKS gene whose product is incorporated into a PKS.
- 15 but also:

10

20

30

- (ii) from fusions of heterologous module (where heterologous module means two modules are adjacent to one another that are not adjacent to one another in naturally occurring PKS enzymes) coding sequences to produce a hybrid coding sequence contained in a PKS gene whose product is incorporated into a PKS,
- (iii) from expression of one or more FK-520 PKS genes with one or more non-FK-520 PKS genes, including both naturally occurring and recombinant non-FK-520 PKS genes, and
- (iv) from combinations of the foregoing.
 Various hybrid PKSs of the invention illustrating these various alternatives are described
 herein.

Examples of the production of a hybrid PKS by co-expression of PKS genes from the FK-520 PKS and another non-FK-520 PKS include hybrid PKS enzymes produced by coexpression of FK-520 and rapamycin PKS genes. Preferably, such hybrid PKS enzymes are produced in recombinant *Streptomyces* host cells that produce FK-520 or FK-506 but have been mutated to inactivate the gene whose function is to be replaced by the rapamycin PKS gene introduced to produce the hybrid PKS. Particular examples include (i) replacement of the *fkbC* gene with the *rapB* gene; and (ii) replacement of the *fkbA* gene with the *rapC* gene. The latter hybrid PKS produces 13,15-didesmethoxy-FK-520, if the host cell

10

15

20

25

30

is an FK-520 producing host cell, and 13,15-didesmethoxy-FK-506, if the host cell is an FK-506 producing host cell. The compounds produced by these hybrid PKS enzymes are immunosuppressants and neurotrophins but can be readily modified to act only as neurotrophins, as described in Example 6, below.

Other illustrative hybrid PKS enzymes of the invention are prepared by replacing the fkbA gene of an FK-520 or FK-506 producing host cell with a hybrid fkbA gene in which: (a) the extender module 8 through 10, inclusive, coding sequences have been replaced by the coding sequences for extender modules 12 to 14, inclusive, of the rapamycin PKS; and (b) the module 8 coding sequences have been replaced by the module 8 coding sequence of the rifan Lin PKS. When expressed with the other, naturally occurring FK-520 or FK-506 PKS genes and the genes of the modification enzymes, the resulting hybrid PKS enzymes produce, respectively, (a) 13-desmethoxy-FK-520 or 13-desmethoxy-FK-506; and (b) 13desmethoxy-13-methyl-FK-520 or 13-desmethoxy-13-methyl-FK-506. In a preferred embodiment, these recombinant PKS genes of the invention are introduced into the producing host cell by a vector such as pHU204, which is a plamsid pRM5 derivative that has the well-characterized SCP2* replicon, the colEl replicon, the tsr and bla resistance genes, and a cos site. This vector can be used to introduce the recombinant fkbA replacement gene in an FK-520 or FK-506 producing host cell (or a host cell derived therefrom in which the endogenous fkbA gene has either been rendered inactive by mutation, deletion or homologous recombination with the gene that replaces it) to produce the desired hybrid PKS.

In constructing hybrid PKSs of the invention, certain general methods may be helpful. For example, it is often beneficial to retain the framework of the module to be altered to make the hybrid PKS. Thus, if one desires to add DH and ER functionalities to a module, it is often preferred to replace the KR domain of the original module with a KR, DH, and ER domain-containing segment from another module, instead of merely inserting DH and ER domains. One can alter the stereochemical specificity of a module by replacement of the KS domain with a KS domain from a module that specifies a different stereochemistry. See Lau et al., 1999, "Dissecting the role of acyltransferase domains of modular polyketide synthases in the choice and stereochemical fate of extender units," *Biochemistry* 38(5):1643-1651, incorporated herein by reference. Stereochemistry can also be changed by changing the KR domain. Also, one can alter the specificity of an AT domain by changing only a small segment of the domain. See Lau et al., supra. One can

also take advantage of known linker regions in PKS proteins to link modules from two different PKSs to create a hybrid PKS. See Gokhale et al., 16 Apr. 1999, "Dissecting and Exploiting Intermodular Communication in Polyketide Synthases," *Science 284*: 482-485, incorporated herein by reference.

The following Table lists references describing illustrative PKS genes and corresponding enzymes that can be utilized in the construction of the recombinant PKSs and the corresponding DNA compounds that encode them of the invention. Also presented are various references describing tailoring enzymes and corresponding genes that can be employed in accordance with the methods of the sent invention.

10 Avermectin

5

U.S. Pat. No. 5.252,474 to Merck.

MacNeil et al., 1993, Industrial Microorganisms: Basic and Applied Molecular Genetics, Baltz, Hegeman, & Skatrud, eds. (ASM), pp. 245-256, A Comparison of the Genes Encoding the Polyketide Synthases for Avermectin, Erythromycin, and Nemadectin.

MacNeil et al., 1992, Gene 115: 119-125, Complex Organization of the Streptomyces avermitilis genes encoding the avermectin polyketide synthase.

Ikeda *et al.*. Aug. 1999, Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*, *Proc. Natl. Acad. Sci. USA* 96: 9509-9514.

20 Candicidin (FR008)

Hu et al., 1994, Mol. Microbiol. 14: 163-172.

Epothilone

U.S. Pat. App. Serial No. 60/130,560, filed 22 April 1999.

Erythromycin

25 PCT Pub. No. 93/13663 to Abbott.

US Pat. No. 5.824,513 to Abbott.

Donadio et al., 1991, Science 252:675-9.

Cortes et al., 8 Nov. 1990, Nature 348:176-8, An unusually large multifunctional polypeptide in the erythromycin producing polyketide synthase of Saccharopolyspora erythraea.

Glycosylation Enzymes

PCT Pat. App. Pub. No. 97/23630 to Abbott.

FK-506

30

Motamedi *et al.*, 1998. The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK-506, *Eur. J. biochem.* 256; 528-534.

Motamedi *et ai.*, 1997, Structurai organization of a multifunctional polyketide synthase involved in the biosynthesis of the macrolide immunosuppressant FK-506, *Eur. J. Biochem. 244*: 74-80.

Methyltransferase

US 5,264,355, issued 23 Nov. 1993, Methylating enzyme from *Streptomyces* MA6858, 31-O-desmethyl-FK-506 methyltransferase.

hydroxylase genes involved in the biosynthesis of the immunosuppressants FK-506 and FK-520, *J. Bacteriol.* 178: 5243-5248.

Motamedi et al., 1996 Characterization of methyltransferase and

Streptomyces hygroscopicus

U.S. patent application Serial No. 09/154,083, filed 16 Sep. 1998.

Lovastatin

U.S. Pat. No. 5,744,350 to Merck.

Narbomycin

U.S. patent application Serial No. 60/107.093, filed 5 Nov. 1998, and Serial No. 60/120,254, filed 16 Feb. 1999.

Nemadectin

20 MacNeil et al., 1993, supra.

Niddamycin

Kakavas *et al.*, 1997. Identification and characterization of the niddamycin polyketide synthase genes from *Streptomyces caelestis*, *J. Bacteriol.* 179: 7515-7522.

Oleandomycin

Swan *et al.*, 1994, Characterisation of a *Streptomyces antibioticus* gene encoding a type I polyketide synthase which has an unusual coding sequence, *Mol. Gen. Genet. 242*: 358-362.

U.S. patent application Serial No. 60/120,254, filed 16 Feb. 1999.

Olano et al., 1998. Analysis of a *Streptomyces antibioticus* chromosomal region involved in oleandomycin biosynthesis, which encodes two glycosyltransferases responsible for glycosylation of the macrolactone ring, *Mol. Gen. Genet.* 259(3): 299-308.

Picromycin

PCT patent application US99/15047, filed 2 Jul. 1999.

Xue *et al.*, 1998, Hydroxylation of macrolactones YC-17 and narbomycin is mediated by the *pik*C-encoded cytochrome P450 in *Streptomyces venezuelae*, *Chemistry & Biology* 5(11): 061-667.

Xue et al., Oct. 1998, A gene cluster for macrolide antibiotic biosynthesis in

Streptomyces venezuelae: Architecture of metabolic diversity, Proc. Natl. Acad. Sci. USA

95: 12111-12116.

Platenolide

EP Pat. App. Pub. No. 791,656 to Lilly.

Papamycin

Schwecke *et al.*, Aug. 1995, The biosynthetic gene cluster for the polyketide rapamycin, *Proc. Natl. Acad. Sci. USA 92:*7839-7843.

Aparicio et al., 1996, Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of the enzymatic domains in the modular polyketide synthase. *Gene 169*: 9-16.

15 Rifamycin

August *et al.*, 13 Feb. 1998, Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the *rif* biosynthetic gene cluster of *Amycolatopsis mediterranei* S669, *Chemistry & Biology*, 5(2): 69-79.

Sorangium PKS

U.S. patent application Serial No. 09/144,085, filed 31 Aug. 1998.

Soraphen

U.S. Pat. No. 5,716,849 to Novartis.

Schupp et al., 1995, J. Bacteriology 177: 3673-3679. A Sorangium cellulosum (Myxobacterium) Gene Cluster for the Biosynthesis of the Macrolide Antibiotic Soraphen

A: Cloning, Characterization, and Homology to Polyketide Synthase Genes from Actinomycetes.

Spiramycin

U.S. Pat. No. 5,098,837 to Lilly.

Activator Gene

30 U.S. Pat. No. 5,514,544 to Lilly.

Tylosin

EP Pub. No. 791,655 to Lilly.

U.S. Pat. No. 5,876,991 to Lilly.

WO 00/20601 PCT/US99/22886

Kuhstoss *et al.*, 1996. *Gene 183*:231-6.. Production of a novel polyketide through the construction of a hybrid polyketide synthase.

Tailoring enzymes

5

10

15

20

25

30

Merson-Davies and Cundliffe, 1994, *Mol. Microbiol.* 13: 349-355. Analysis of five tylosin biosynthetic genes from the *tylBA* region of the *Streptomyces fradiae* genome.

As the above Table illustrates, there are a wide variety of polyketide synthase genes that serve as readily available sources of DNA and sequence information for use in constructing the hybrid PKS-encoding DNA compounds of the invention. Methods for constructing hybrid PKS-encoding DNA compounds are described without reference to the FK-520 PKS in PCT patent publication No. 98-51695; U.S. Patent Nos. 5,672,491 and 5,712,146 and U.S. patent application Serial Nos. 09/073.538, filed 6 May 1998, and 09/141,908, filed 28 Aug 1998, each of which is incorporated herein by reference.

The hybrid PKS-encoding DNA compounds of the invention can be and often are hybrids of more than two PKS genes. Moreover, there are often two or more modules in the hybrid PKS in which all or part of the module is derived from a second (or third) PKS. Thus, as one illustrative example, the present invention provides a hybrid FK-520 PKS that contains the naturally occurring loading module and FkbP as well as modules one, two, four, six, seven, and eight, nine, and ten of the FK-520 PKS and further contains hybrid or heterologous modules three and five. Hybrid or heterologous module three contains an AT domain that is specific of methylmalonyl CoA and can be derived for example, from the erythromycin or rapamycin PKS genes. Hybrid or heterologous module five contains an AT domain that is specific for malonyl CoA and can be derived for example, from the picromycin or rapamycin PKS genes.

While an important embodiment of the present invention relates to hybrid PKS enzymes and corresponding genes, the present invention also provides recombinant FK-520 PKS genes in which there is no second PKS gene sequence present but which differ from the FK-520 PKS gene by one or more deletions. The deletions can encompass one or more modules and/or can be limited to a partial deletion within one or more modules. When a deletion encompasses an entire module, the resulting FK-520 derivative is at least two carbons shorter than the gene from which it was derived. When a deletion is within a module, the deletion typically encompasses a KR, DH, or ER domain, or both DH and ER domains, or both KR and DH domains, or all three KR, DH, and ER domains.

15

20

25

30

To construct a hybrid PKS or FK-520 derivative PKS gene of the invention, one can employ a technique, described in PCT Pub. No. 98/27203 and U.S. patent application Serial No. 08 989,332. Mied 11 Dec. 1997, each of which is incorporated herein by reference, in which the large PKS gene is divided into two or more, typically three, segments, and each segment is placed on a separate expression vector. In this manner, each of the segments of the gene can be altered, and various altered segments can be combined in a single host cell to provide a recombinant PKS gene of the invention. This technique makes more efficient the construction of large libraries of recombinant PKS genes, vectors for expressing those genes, and host cells comprising those vectors.

Thus, in one important embodiment, the recombinant DNA compounds of the invention are expression vectors. As used herein, the term expression vector refers to any nucleic acid that can be introduced into a host cell or cell-free transcription and translation medium. An expression vector can be maintained stably or transiently in a cell, whether as part of the chromosomal or other DNA in the cell or in any cellular compartment, such as a replicating vector in the cytoplasm. An expression vector also comprises a gene that serves to produce RNA that is translated into a polypeptide in the cell or cell extract. Furthermore, expression vectors typically contain additional functional elements, such as resistance-conferring genes to act as selectable markers.

The various components of an expression vector can vary widely, depending on the intended use of the vector. In particular, the components depend on the host cell(s) in which the vector will be used or is intended to function. Vector components for expression and maintenance of vectors in *E. coli* are widely known and commercially available, as are vector components for other commonly used organisms, such as yeast cells and *Streptomyces* cells.

In a preferred embodiment, the expression vectors of the invention are used to construct recombinant *Streptomyces* host cells that express a recombinant PKS of the invention. Preferred *Streptomyces* host cell/vector combinations of the invention include *S. coelicolor* CH999 and *S. lividans* K4-114 host cells, which do not produce actinorhodin, and expression vectors derived from the pRM1 and pRM5 vectors, as described in U.S. Patent No. 5,830,750 and U.S. patent application Serial Nos. 08/828,898, filed 31 Mar. 1997, and 09/181,833, filed 28 Oct. 1998, each of which is incorporated herein by reference.

25

30

The present invention provides a wide variety of expression vectors for use in Streptomyces. For replicating vectors, the origin of replication can be, for example and without limitation, a low copy number vector, such as SCP2* (see Hopwood et al., Genetic Manipulation of Streptomyces: A Laboratory manual (The John Innes Foundation. 5 Norwich, U.K., 1985); Lydiate et al., 1985, Gene 35: 223-235; and Kieser and Melton, 1988, Gene 65: 83-91, each of which is incorporated herein by reference), SLP1.2 (Thompson et al., 1982, Gene 20: 51-62, incorporated herein by reference), and SG5(ts) (Muth et al., 1989, Mol. Gen. Genet. 219: 341-348, and Bierman et al., 1992, Gene 116: 43-49, each of which is incorporated herein by reference), or a high copy number vector, such 10 as pIJ101 and pJV1 (see Katz et al., 1983, J. Gen. Microbio! 29: 2703-2714; Var., et al., 1989, J. Bacteriol. 171: 5782-5781; and Servin-Gonzalez. 1993, Plasmid 30: 131-140, each of which is incorporated herein by reference). Generally, however, high copy number vectors are not preferred for expression of genes contained on large segments of DNA. For non-replicating and integrating vectors, it is useful to include at least an E. coli origin of 15 replication, such as from pUC, p1P, p1I, and pBR. For phage based vectors, the phages phiC31 and KC515 can be employed (see Hopwood et al., supra).

Typically, the expression vector will comprise one or more marker genes by which host cells containing the vector can be identified and/or selected. Useful antibiotic resistance conferring genes for use in *Streptomyces* host cells include the *ermE* (confers resistance to erythromycin and other macrolides and lincomycin). *tsr* (confers resistance to thiostrepton), *aadA* (confers resistance to spectinomycin and streptomycin), *aacC4* (confers resistance to apramycin, kanamycin, gentamicin, geneticin (G418), and neomycin), *hyg* (confers resistance to hygromycin), and *vph* (confers resistance to viomycin) resistance conferring genes.

The recombinant PKS gene on the vector will be under the control of a promoter, typically with an attendant ribosome binding site sequence. The present invention provides the endogenous promoters of the FK-520 PKS and related biosynthetic genes in recombinant form, and these promoters are preferred for use in the native hosts and in heterologous hosts in which the promoters function. A preferred promoter of the invention is the fkbO gene promoter, comprised in a sequence of about 270 bp between the start of the open reading frames of the fkbO and fkbB genes. The fkbO promoter is believed to be bidirectional in that it promotes transcription of the genes fkbO, fkbP, and fkbA in one direction and fkbB, fkbC, and fkbL in the other. Thus, in one aspect, the present invention

10

15

20

25

30

provides a recombinant expression vector comprising the promoter of the *fkbO* gene of an FK-520 producing organism positioned to transcribe a gene other than *fkbO*. In a preferred embodiment the transcribed gene is an FK-520 PKS gene. In another preferred embodiment, the transcribed gene is a gene that encodes a protein comprised in a hybrid PKS.

Heterologous promoters can also be employed and are preferred for use in host cells in which the endogenous FK-520 PKS gene promoters do not function or function poorly. A preferred heterologous promoter is the actI promoter and its attendant activator gene actII-ORF4, which is provided in the pRM1 and pRM5 expression vectors, supra. This promoter is activated but stationary phase of growth when secondary metabolites are normalisynthesized. Other useful Streptomyces promoters include without limitation those from the ermE gene and the melC1 gene, which act constitutively, and the tipA gene and the merA gene, which can be induced at any growth stage. In addition, the T7 RNA polymerase system has been transferred to Streptomyces and can be employed in the vectors and host cells of the invention. In this system, the coding sequence for the T7 RNA polymerase is inserted into a neutral site of the chromosome or in a vector under the control of the inducible merA promoter, and the gene of interest is placed under the control of the T7 promoter. As noted above, one or more activator genes can also be employed to enhance the activity of a promoter. Activator genes in addition to the actII-ORF4 gene discussed above include dnr1, redD, and ptpA genes (see U.S. patent application Serial No. 09/181,833, supra) to activate promoters under their control.

In addition to providing recombinant DNA compounds that encode the FK-520 PKS, the present invention also provides DNA compounds that encode the ethylmalonyl CoA and 2-hydroxymalonyl CoA utilized in the synthesis of FK-520. Thus, the present invention also provides recombinant host cells that express the genes required for the biosynthesis of ethylmalonyl CoA and 2-hydroxymalonyl CoA. Figures 3 and 4 show the location of these genes on the cosmids of the invention and the biosynthetic pathway that produces ethylmalonyl CoA.

For 2-hydroxymalonyl CoA biosynthesis, the *fkbH*, *fkbI*, *fkbJ*, and *fkbK* genes are sufficient to confer this ability on *Streptomcyces* host cells. For conversion of 2-hydroxymalonyl to 2-methoxymalonyl, the *fkbG* gene is also employed. While the complete coding sequence for *fkbH* is provided on the cosmids of the invention, the sequence for this gene provided herein may be missing a T residue, based on a comparison made with a similar gene cloned from the ansamitocin gene cluster by Dr. H. Floss. Where the sequence

15

20

25

30

WO 00/20601 PCT/US99/22886

herein shows one T, there may be two, resulting in an extension of the *fkbH* reading frame to encode the amino acid sequence:

MTIVKCLVWDLDNTLWRGTVLEDDEVVLTDEIREVITTLDDRGILQAVASKNDHD LAWERLERLGVAEYFVLARIGWGPKSQSVREIATELNFAPTTIAFIDDQPAERAEVA FHLPEVRCYPAEQAATLLSLPEFSPPVSTVDSRRRRLMYQAGFARDQAREAYSGPD EDFLRSLDLSMTIAPAGEEELSRVEELTLRTSQMNATGVHYSDADLRALLTDPAHE VLVVTMGDRFGPHGAVGIILLEKKPSTWHLKLLATSCRVVSFGAGATILNWLTDQG ARAGAHLVADFRRTDRNRMMEIAYRFAGFADSDCPCVSEVAGASAAGVERLHLEP SARPAPTTLTLTAADIAPVTVSAAG.

For ethylmalonyl CoA biosynthesis, one requires only a crotonyl CoA reductase, which can be supplied by the host cell but can also be supplied by recombinant expression of the *fkbS* gene of the present invention. To increase yield of ethylmalonyl CoA, one can also express the *fkbE* and *fkbU* genes as well. While such production can be achieved using only the recombinant genes above, one can also achieve such production by placing into the recombinant host cell a large segment of the DNA provided by the cosmids of the invention. Thus, for 2-hydroxymalonyl and 2-methoxymalonyl CoA biosynthesis, one can simply provide the cells with the segment of DNA located on the left side of the FK-520 PKS genes shown in Figure 1. For ethylmalonyl CoA biosynthesis, one can simply provide the cells with the segment of DNA located on the right side of the FK-520 PKS genes shown in Figure 1 or, alternatively, both the right and left segments of DNA.

The recombinant DNA expression vectors that encode these genes can be used to construct recombinant host cells that can make these important polyketide building blocks from cells that otherwise are unable to produce them. For example, *Streptomyces coelicolor* and *Streptomyces lividans* do not synthesisze ethylmalonyl CoA or 2-hydroxymalonyl CoA. The invention provides methods and vectors for constructing recombinant *Streptomyces*

coelicolor and Streptomyces lividans that are able to synthesize either or both ethylmalonyl CoA and 2-hydroxymalonyl CoA. These host cells are thus able to make polyketides, those requiring these substrates, that cannot otherwise be made in such cells.

In a preferred embodiment, the present invention provides recombinant Streptomyces host cells, such as S. coelicolor and S. lividans, that have been transformed with a recombinant vector of the invention that codes for the expression of the ethylmalonyl CoA biosynthetic genes. The resulting host cells produce ethylmalonyl CoA and so are preferred host cells for the production of polyketides produced by PKS enzymes that

15

20

25

30

WO 00/20601 PCT/US99/22886

comprise one or more AT domains specific for ethylmalonyl CoA. Illustrative PKS enzymes of this type include the FK-520 PKS and a recombinant PKS in which one or more AT domains is specific for ethylmalonyl CoA.

In a related embodiment, the present invention provides *Streptomyces* host cells in which one or more of the ethylmalonyl or 2-hydroxymalonyl biosynthetic genes have been deleted by homologous recombination or rendered inactive by mutation. For example, deletion or inactivation of the *fkbG* gene can prevent formation of the methoxyl groups at C-13 and C-15 of FK-520 (or, in the corresponding FK-506 producing cell, FK-506), leading to the production of 13,15-didesmethoxy-13,15-dihydroxy-FK-520 (or, in the corresponding FK-506 producing cell, 13,15-didesmethoxy-13,15-dihydroxy-FK-506). If the *fkbG* gene product acts on 2-hydroxymalonyl and the resulting 2-methoxymalonyl substrate is required for incorporation by the PKS, the AT domains of modules 7 and 8 may bind malonyl CoA and methylmalonyl CoA. Such incorporation results in the production of a mixture of polyketides in which the methoxy groups at C-13 and C-15 of FK-520 (or FK-506) are replaced by either hydrogen or methyl.

This possibility of non-specific binding results from the construction of a hybrid PKS of the invention in which the AT domain of module 8 of the FK-520 PKS replaced the AT domain of module 6 of DEBS. The resulting PKS produced, in *Streptomyces lividans*, 6-dEB and 2-desmethyl-6-dEB, indicating that the AT domain of module 8 of the FK-520 PKS could bind malonyl CoA and methylmalonyl CoA substrates. Thus, one could possibly also prepare the 13.15-didesmethoxy-FK-520 and corresponding FK-506 compounds of the invention by deleting or otherwise inactivating one or more or all of the genes required for 2-hydroxymalonyl CoA biosynthesis. i.e., the *fkbH*, *fkbI*, *fkbJ*, and *fkbK* genes. In any event, the deletion or inactivation of one or more biosynthetic genes required for ethylmalonyl and/or 2-hydroxymalonyl production prevents the formation of polyketides requiring ethylmalonyl and/or 2-hydroxymalonyl for biosynthesis, and the resulting host cells are thus preferred for production of polyketides that do not require the same.

The host cells of the invention can be grown and fermented under conditions known in the art for other purposes to produce the compounds of the invention. See, e.g., U.S. Patent Nos. 5,194,378; 5,116,756; and 5,494,820, incorporated herein by reference, for suitable fermentation processes. The compounds of the invention can be isolated from the fermentation broths of these cultured cells and purified by standard procedures. Preferred compounds of the invention include the following compounds: 13-desmethoxy-FK-506; 13-

10

15

20

25

30

desmethoxy-FK-520; 13,15-didesmethoxy-FK-506; 13,15-didesmethoxy-FK-520; 13-desmethoxy-18-hydroxy-FK-506; 13-desmethoxy-18-hydroxy-FK-520; 13,15-didesmethoxy-18-hydroxy-FK-500; and 13,15-didesmethoxy-18-hydroxy-FK-520. These compounds can be further modified as described for tacrolimus and FK-520 in U.S. Patent Nos. 5,225,403; 5,189,042; 5,164,495; 5,068,323; 4,980,400; and 4,920,218, incorporated herein by reference.

Other compounds of the invention are shown in Figure 8. Part A and B. In Figure 8. Part A, illustrative C-32-substituted compounds of the invention are shown in two columns under the heading R. The substituted compounds are preferred for topical administration and are applied to the dermis for treatment — conditions such as psoriasis. In Figure 8, Part B, illustrative reaction schemes for making the compounds shown in Figure 8, Part A, are provided. In the upper scheme in Figure 8, Part B, the C-32 substitution is a tetrazole moiety, illustrative of the groups shown in the left column under R in Figure 8, Part A. In the lower scheme in Figure 8, Part B, the C-32 substitution is a disubstituted amino group, where R₃ and R₄ can be any group similar to the illustrative groups shown attached to the amine in the right column under R in Figure 8. Part A. While Figure 8 shows the C-32-substituted compounds in which the C-15-methoxy is present, the invention includes these C-32-substituted compounds in which C-15 is ethyl, methyl, or hydrogen. Also, while C-21 is shown as substituted with ethyl or allyl, the compounds of the invention includes the C-32-substituted compounds in which C-21 is substituted with hydrogen or methyl.

To make these C-32-substituted compounds, Figure 8, Part B, provides illustrative reaction schemes. Thus, a selective reaction of the starting compound (see Figure 8, Part B, for an illustrative starting compound) with trifluoromethanesulfonic anhydride in the presence of a base yields the C-32 O-triflate derivative, as shown in the upper scheme of Figure 8, Part B. Displacement of the triflate with 1H-tetrazole or triazole derivatives provides the C-32 tetrazole or teiazole derivative. As shown in the lower scheme of Figure 8, Part B, reacting the starting compound with p-nitrophenylchloroformate yields the corresponding carbonate, which, upon displacement with an amino compound, provides the corresponding carbamate derivative.

The compounds can be readily formulated to provide the pharmaceutical compositions of the invention. The pharmaceutical compositions of the invention can be used in the form of a pharmaceutical preparation, for example, in solid, semisolid, or liquid form. This preparation contains one or more of the compounds of the invention as an active

10

15

20

25

30

ingredient in admixture with an organic or inorganic carrier or excipient suitable for external, enteral, or parenteral application. The active ingredient may be compounded, for example, with the usual non-toxic, pharmaceutically acceptable carriers for tablets, pellets, capsules, suppositories, solutions, emulsions, suspensions, and any other form suitable for use. Suitable formulation processes and compositions for the compounds of the present invention are described with respect to tacrolimus in U.S. Patent Nos. 5,939,427; 5,922,729; 5,385,907; 5,338,684; and 5,260,301, incorporated herein by reference. Many of the compounds of the invention contain one or more chiral centers, and all of the stereoisomers are included within the scope of the invention, as pure compounds as well as mixtures of stereoisomers. Thus the compounds of the invention may be supplied as a mixture of stereoisomers in any proportion.

The carriers which can be used include water, glucose, lactose, gum acacia, gelatin, mannitol, starch paste, magnesium trisilicate, tale, corn starch, keratin, colloidal silica, potato starch, urea, and other carriers suitable for use in manufacturing preparations, in solid, semi-solid, or liquified form. In addition, auxiliary stabilizing, thickening, and coloring agents and perfumes may be used. For example, the compounds of the invention may be utilized with hydroxypropyl methylcellulose essentially as described in U.S. Patent No. 4,916,138, incorporated herein by reference, or with a surfactant essentially as described in EPO patent publication No. 428,169, incorporated herein by reference.

Oral dosage forms may be prepared essentially as described by Hondo *et al.*, 1987, *Transplantation Proceedings XIX*, Supp. 6: 17-22, incorporated herein by reference. Dosage forms for external application may be prepared essentially as described in EPO patent publication No. 423,714, incorporated herein by reference. The active compound is included in the pharmaceutical composition in an amount sufficient to produce the desired effect upon the disease process or condition.

For the treatment of conditions and diseases relating to immunosuppression or neuronal damage, a compound of the invention may be administered orally, topically, parenterally, by inhalation spray, or rectally in dosage unit formulations containing conventional non-toxic pharmaceutically acceptable carriers, adjuvant, and vehicles. The term parenteral, as used herein, includes subcutaneous injections, and intravenous, intramuscular, and intrasternal injection or infusion techniques.

Dosage levels of the compounds of the present invention are of the order from about 0.01 mg to about 50 mg per kilogram of body weight per day, preferably from about 0.1 mg

15

20

25

30

to about 10 mg per kilogram of body weight per day. The dosage levels are useful in the treatment of the above-indicated conditions (from about 0.7 mg to about 3.5 mg per patient per day, assuming a 70 kg patient). In addition, the compounds of the present invention may be administered on an intermittent basis, i.e., at semi-weekly, weekly, semi-monthly, or monthly intervals.

The amount of active ingredient that may be combined with the carrier materials to produce a single dosage form will vary depending upon the host treated and the particular mode of administration. For example, a formulation intended for oral administration to humans may contain from 0.5 mg to 5 g of active agent compounded with an appropriate and convenient amount of carrier material, which may vary from about 5 percent to about 95 percent of the total composition. Dosage unit forms will generally contain from about 0.5 mg to about 500 mg of active ingredient. For external administration, the compounds of the invention can be formulated within the range of, for example, 0.00001% to 60% by weight, preferably from 0.001% to 10% by weight, and most preferably from about 0.005% to 0.8% by weight. The compounds and compositions of the invention are useful in treating disease conditions using doses and administration schedules as described for tacrolimus in U.S. Patent Nos. 5,542,436; 5,365,948; 5,348,966; and 5,196,437, incorporated herein by reference. The compounds of the invention can be used as single therapeutic agents or in combination with other therapeutic agents. Drugs that can be usefully combined with compounds of the invention include one or more immunosuppressant agents such as rapamycin, cyclosporin A, FK-506, or one or more neurotrophic agents.

It will be understood, however, that the specific dosage level for any particular patient will depend on a variety of factors. These factors include the activity of the specific compound employed; the age, body weight, general health, sex. and diet of the subject; the time and route of administration and the rate of excretion of the drug; whether a drug combination is employed in the treatment; and the severity of the particular disease or condition for which therapy is sought.

A detailed description of the invention having been provided above, the following examples are given for the purpose of illustrating the present invention and shall not be construed as being a limitation on the scope of the invention or claims.

Example 1

Replacement of Methoxvl with Hvdrogen or Methyl at C-13 of FK-520

15

20

25

30

The C-13 methoxyl group is introduced into FK-520 via an AT domain in extender module 8 of the PKS that is specific for hydroxymalonyl and by methylation of the hydroxyl group by an S-adenosyl methionine (SAM) dependent methyltransferase. Metabolism of FK-506 and FK-520 primarily involves oxidation at the C-13 position into an inactive derivative that is further degraded by host P450 and other enzymes. The present invention provides compounds related in structure to FK-506 and FK-520 that do not contain the C-13 methoxy group and exhibit greater stability and a longer half-life *in vivo*. These compounds are useful medicaments due to their immunosuppressive and neurotrophic activities, and the invention provides the compounds in purified form and as pharmaceutical compositions.

The present invention also provides the novel PKS enzymes that produce these novel compounds as well as the expression vectors and host cells that produce the novel PKS enzymes. The novel PKS enzymes include, among others, those that contain an AT domain specific for either malonyl CoA or methylmalonyl CoA in module 8 of the FK-506 and FK-520 PKS. This example describes the construction of recombinant DNA compounds that encode the novel FK-520 PKS enzymes and the transformation of host cells with those recombinant DNA compounds to produce the novel PKS enzymes and the polyketides produced thereby.

To construct an expression cassette for performing module 8 AT domain replacements in the FK-520 PKS, a 4.6 kb SphI fragment from the FK-520 gene cluster was cloned into plasmid pLitmus 38 (a cloning vector available from New England Biolabs). The 4.6 kb SphI fragment, which encodes the ACP domain of module 7 followed by module 8 through the KR domain, was isolated from an agarose gel after digesting the cosmid pKOS65-C31 with Sph I. The clone having the insert oriented so the single SacI site was nearest to the SpeI end of the polylinker was identified and designated as plasmid pKOS60-21-67. To generate appropriate cloning sites, two linkers were ligated sequentially as follows. First, a linker was ligated between the SpeI and SacI sites to introduce a BglII site at the 5' end of the cassette, to eliminate interfering polylinker sites, and to reduce the total insert size to 4.5 kb (the limit of the phage KC515). The ligation reactions contained 5 picomolar unphosphorylated linker DNA and 0.1 picomolar vector DNA, i.e., a 50-fold molar excess of linker to vector. The linker had the following sequence:

5'-CTAGTGGGCAGATCTGGCAGCT-3' 3'-ACCCGTCTAGACCG-5'

The resulting plasmid was designated pKOS60-27-1.

10

25

Next, a linker of the following sequence was ligated between the unique *SphI* and *AfIII* sites of plasmid pKOS60-27-1 to introduce an *NsiI* site at the 3' end of the module 8 cassette. The linker employed was:

5'-GGGATGCATGGC-3'
3'-GTACCCCTACGTACCGAATT-5'

The resulting plasmid was designated pKOS60-29-55.

To allow in-frame insertions of alternative AT domains, sites were engineered at the 5' end (Avr II or Nhe I) and 3' end (Aho I) of the AT domain using the polymerase chain reaction (PCR) as follows. Plasmid pKOS60-29-55 was used as a template for the PCR and sequence 5' to the AT domain was amplified with the primers SpeBgl-fwd and either Avr-rey or Nhe-rey.

SpeBgl-fwd 5'-CGACTCACTAGTGGGCAGATCTGG-3'

Avr-rev 5'-CACGCCTAGGCCGGTCGGTCTCGGGCCAC-3'

Nhe-rev 5'-GCGGCTAGCTGCTCGCCCATCGCGGGATGC-3'

The PCR included, in a 50 μl reaction, 5 μl of 10x Pfu polymerase buffer (Stratagene), 5 μl 10x z-dNTP mixture (2 mM dATP, 2 mM dCTP, 2 mM dTTP, 1 mM dGTP, 1 mM 7-deaza-GTP), 5 μl DMSO, 2 μl of each primer (10 μM), 1 μl of template DNA (0.1 μg/μl), and 1 μl of cloned Pfu polymerase (Stratagene). The PCR conditions were 95°C for 2 min., 25 cycles at 95°C for 30 sec., 60°C for 30 sec., and 72°C for 4 min., followed by 4 min. at 72°C and a hold at 0°C. The amplified DNA products and the Litmus vectors were cut with the appropriate restriction enzymes (Bg/II and AvrII or SpeI and NheI), and cloned into either pLitmus 28 or pLitmus38 (New England Biolabs), respectively, to generate the constructs designated pKOS60-37-4 and pKOS60-37-2, respectively.

Plasmid pKOS60-29-55 was again used as a template for PCR to amplify sequence 3' to the AT domain using the primers BsrXho-fwd and NsiAfl-rev:

BsrXho-fiwd 5'-GATGTACAGCTCGAGTCGGCACGCCCGGCCGCATC-3'
NsiAfl-rev 5'-CGACTCACTTAAGCCATGCATCC-3'

PCR conditions were as described above. The PCR fragment was cut with *BsrGI* and *AfIII*, gel isolated, and ligated into pKOS60-37-4 cut with *Asp*718 and *AfIII* and inserted into pKOS60-37-2 cut with *BsrGI* and *AfIII*, to give the plasmids pKOS60-39-1 and pKOS60-39-13, respectively. These two plasmids can be digested with *AvrII* and *XhoI* or *NheI* and *XhoI*, respectively, to insert heterologous AT domains specific for malonyl, methylmalonyl, ethylmalonyl, or other extender units.

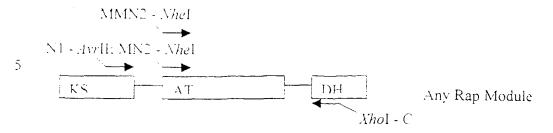
Malonyl and methylmalonyl-specific AT domains were cloned from the rapamycin cluster using PCR amplification with a pair of primers that introduce an *Avr*II or *Nhe*I site at the 5° end and an *Xho*I site at the 3° end. The PCR conditions were as given above and the primer sequences were as follows:

5

RATN1 5'-ATCCTAGGCGGGCRGGYGTGTCGTCCTTCGG-3' (3' end of Rap KS sequence and universal for malonyl and methylmalonyl CoA), RATMN2 5'-ATGCTAGCCGCCGCGTTCCCCGTCTTCGCGCG-3' (Rap AT shorter version 5'- sequence and specific for malonyl CoA),

10 RATMMN2 5'-ATGCTAGCGGATTCGTCGGTGGTGTTCGCCGA-3' (Rap AT shorter version 5'- sequence and specific for methylmalonyl CoA), and RATC 5'-ATCTCGAGCCAGTASCGCTGGTGYTGGAAGG-3' (Rap DH 5'- sequence and universal for malonyl and methylmalonyl CoA).

15



Because of the high sequence similarity in each module of the rapamycin cluster, each primer was expected to prime any of the AT domains. PCR products representing ATs specific for malonyl or methylmalonyl extenders were identified by sequencing individual cloned "TR products. Sequencing also confirmed that the chosen clones contained no cloning artifacts. Examples of hybrid modules with the rapamycin AT12 and AT13 domains are shown in a separate figure.

The AvrII-AhoI restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 12 of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below. The AT of rap module 12 is specific for incorporation of malonyl units.

20 AGATOTGGCAGOTCGCCGAAGOGCTGCTGACGCTCGTCCGGGAGAGCACC 50 I W O L A E A L L T L V R E S T GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGGCGGC 100 A A V L G H V G G E D I P A T A A GTTCAAGGACCTCGGCATCGACTCGCTCACCGCGGTCCAGCTGCGCAACG 150 25 F K D L G I D S L T A V Q L R N COCTCACOGAGGCGACOGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200 A L T E A T G V R L N A T A V F D TTOCCSRCCCCGCACSTGCTCGCCGGGARGCTCGGCGACGAACTGACCGG 250 FPTPHULAGKLGPELTG 30 CACCCGCGCGCCCCTCGTGCCCCGGACCGCGCCCCCACGCCGGTGCGGTACG 300 T R A P V V P R T A A T A G A H ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGGCTGCCCGGCGGGGGGGTC 350 D E P L A I V G M A C E L P G G V GCGTCACCCGAGGASCTSTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400 35 A S P E E L W H L V A S G T D A CACGGAGTTCCCGACGGACGGCGGCTGGGACGTCGACGCGATCTACGACC 450 TEFFTDRGWDVDAIYD CGGACCCCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500 P D P D A I G K T F V R H G G F L 40 ACCGGCGCGACAGGCTTCGACGCGGCGTTCTTCGGCATCAGCCCGCGCGA 550 TGATGFDAAFFGISPRE GGCCCTCGCGATGGACCCGCAGCAGCGGGTGCTCCTGGAGACGTCGTGGG 600 A L A M D P Q Q R V L L E T S W AGGCGTTCGAAAGCGCCGGCATCACCCCGGACTCGACCCGCGGCAGCGAC 650 45 E A F E S A G I T P D S T R G S D ACCGCCGTGTTCGTCGGCCCTTCTCCTACGGTTACGGCACCGGTGCGGA 700 T G V F Y S A F S Y G Y G T S A D DACCGACGGCTTCGGCGCGACCGGCTCGCAGACCAGTGTGCTCTCCGGCC 750 T D G F G A T G S Q T S V L S G 50 GGCTGTCGTACTTCTACGGTCTGGAGGGTCCGGCGGTCACGGTCGACACG 800

	- 90 STATT TOTO STORED BOT SERVICES TO BOARD AS FOR SOME DAY FOR THE BOT OF SOME STORED BOT SO	3 88.
5		: 1
-	a de la compacta de l Compacta de la compacta de la	1351
	- U F 3 G F M D F 3 R [] F G G L A F D - 660099903AAGSCCTTCGGCGCGCGTTCGGGACCTTCCCCGA	1 100
10	- 3 A A M A B G A B A B G T S F A E - 33GTGCCGGTGTGTGGAGAGGCTCTCCGAGGGCGAACGCAACG	: 155
	GASVLIVERLSCAEFN	
	GTCACACCGTCCTGGCGGTCGTCCGTGGTTCGGCGGTCAACCAGGATGGT G H T T L A V V E G S A V N Q D G	1100
15	- JOSTOMAADAGGITATOAGSWOOGAADAGGOUTTOACAGGAGGAGGATAAT - A 8 M G L 8 A P M G P 8 Q E E 7 7 1	1150
	oddebabu — W. Ummodabodbodadagongodbabeshow A d i A a a g b b a d A i A c A	1200
	TRUMAGGOORACGGCACCGGCACCAGGCTGGGCGGACCCCATCGAGGCACAG	1250
20	GOGGTACTOGOCÁCCTACGGACAGGAGCAGCCACCCTCCTGCTGCT	1300
	A MILIA OT MISIQUE BIA DI PILLULI SI SCONDERDE SE SE TO SE COACO SE CACO DE COMO D	1361
25	3 1 1 7 M V Q A L B H G E L P P T	2.450
	TUGCACGCCGACGAGCCGTCGCCGCCCGTCGACTGGACGGCCGGC	1450
	- OGAACTSCTGACGTCGGCCCGGGCCTAGGC - E L I T S A R P W P E T D R P E	1500
30	GGGCAGGCGTGTCGTCGGGATCAGTGCCACCAACGCCCACGTCATC	1550
30	R A G V S S F G I S G T N A H V I CTGGAAAGGGCACCCCCACTCAGCCTGCGGACAACGCGGTGATCGAGGG	1550
	R A G V S S F G I S G T N A H V I	
30 35	R A G V S S F G I S G T N A H V I CTGGAAAGCGCACCCCCACTCAGCCTGCGGACAACGCGGTGATCGAGCG L E S A F F T Q F A D N A V I E R	1600
	R A G V S S F G I S G T N A H V I CTGGAAAGGCCCCCCCACTCAGCCTGCGGACAACGCGGTGATCGAGCG L E S A F F T Q F A D N A V I E R GGCACCGGACTGGGTGATTTCGGCCAGGACCCAGTCGGCTT A F E W V F L V I S A R T Q S A TGACTGAGCAGAGGCCGGTTGCGTGCTTATCTGGCGGGCG	1600 1650
35	R A G V S S F G I S G T N A H V I CTGGAAAGGCACCCCCACTCAGCCTGCGGACAACGCGGTGATCGAGCG L E S A F F T Q F A D N A V I E R GGCACCGGACTGGGTGATTTCGGCCAGGACCCAGTCGGCTT A F E N V F L V I S A R T Q S A CGACTGAGCACGAGGCCCGGTTGGGTGCGTATCTGGGGGGCGTCGCCCGG L T E H E G R L R A Y L A A S P G GTGGATATCCGGGGTTTTTTGGGTGGGTGACACGGTGTTT V C M F A V A S T L A M T A S V F	1600 1650 1700 1750
	R A G V S S F G I S G T N A H V I CTGGAAAGGGCACCCCCACTCAGCCTGCGGACAACGCGGTGATCGAGCG L E S A F F T Q F A D N A V I E R GGCACCGGAGTGCGTTGGTGATTTCGGCCAGGACCCAGTCGGCTT A F E W V F L V I S A R T Q S A TGACTGAGCACGAGGCCGGTTGGGTGCGTATCTGGCGGGGTTGGCCGGG L T E H E G F L R A Y L A A S F G CTGGATATCCGGGGTTGTTGGGTGGGTGGCGACACGGTCGCTTT V C M E A V A S T L A N T A S V F CGAGCACCGTGCGTGCTGGGGGAACCGGCACCGGTG E H F A V L L G C D T V T G T A	1650 1700 1700 1700
35	R A G V S S F G I S G T N A H V I CTGGAAAGGCACCCCCACTCAGCCTGCGGACAACGCGGTGATCGAGCG L E S A F F T Q F A D N A V I E R GGCACCGGAGTGCGTTGGTGATTTCGGCCAGGACCCAGTCGGCTT A F E M V F L V I S A R T Q S A TGACTGAGCACGAGGGCCGGTTGGGTGGTTATCTGGCGGGGGTCGCCCGGG L T E H E G R L R A Y L A A S F G GTGGATATCCGGGGTTGGGTGGTTGGTTGCGTGGTGTT V C M E A V A S T L A M T F R V F CGAGGACCCTGCGTGGTGGTGACACCGGCACCGGTT E H F A V L L G D D T V T G T A TGTCTGACTCTGGGCGGTGTTTCGTCTTCCCGGGACAGGGTCGCAGGGT	1600 1650 1700 1750
35 40	R A G V S S F G I S G T N A H V I CTGGAAAGGCACCCCCACTCAGCCTGCGGACAACGCGGTGATCGAGCG L E S A F F T Q F A D N A V I E R GGCACCGGACTGGGTGATTTCGGCCAGGACCCAGTCGGCTT A F E N V F L V I S A R T Q S A TGACTGAGCACGAGGGCCGGTTGGGTGGTTATCTGGGGGGGCTCGCCCGGG L T E H E G R L R A Y L A A S F G GTGGATATCCGGGTGCTGCTTGGGTGGGTATCTGGGGGGCTCGCCCGGG V C N F A V L S T L A N T A S V F CGAGCACCCTGCGTGCTGCTGGGGAACCCGGTCACCGGCACCGCTT E H F A V L L G D D T V T G T A TGTCTGACTCTGGGGGGCGGTCTCCCGGGGACACGGGTCCCCGGTCT V S D F R A V F V F P G Q G S Q R GCTGGGCATCGTGGGAACTGGCCGCCGGTCTCCCCGGGACAC	1650 1700 1700 1700 1800
35	R A G V S S F G I S G T N A H V I CTGGAAAGGCACCCCCACTCAGCCTGCGGACAACGCGGTGATCGAGCG L E S A F F T Q F A D N A V I E R GGCACCGGAGTGGGTGCGTTTCGGCCAGGACCCAGTCGGCTT A F E N V F L V I S A R T Q S A CGACTGAGCACGAGGCCGGGTTGGGTGCGTATCTGGGGGGGCGTGGCCGGG L T E H E G R L R A Y L A A S F G GTGGATATCCGGGGTGCTGGTGGGTGACACGGTGGTGTT V C M F A V A S T L A M T F S V F CGAGCACCGTGCGTGCTGTGTGTCACCGGGACCGGTG E H F A V L L G C D T V T G T A TGTCTGACTCTCGGGGGGGGTGTCCCGGGGACACGCT V S C F R A W F V F P G Q G S Q R GCCATCAGCAGGTGTGGGACCGGCGGGGGGGGGACCGGGACACGGTGACCGGGGACACGGTGACCGGGGACACGGTGACCGGGGACACGGTGACCGGGGACACGGTGACCGGGGACACGGTGACCGGGGGGTGGCGGGGGTGGCGGGGGGGG	1600 1650 1700 1700 1800 1800
35 40	R A G V S S F G I S G T N A H V I CTGGAAAGGCACCCCCACTCAGCCTGCGGACAACGCGGTGATCGAGCG L E S A F F T Q F A D N A V I E R GGCACCGGAGTGGGTGCGTTGGGTGATTTCGGCCAGGACCCAGTCGGCTT A F E W V F L V I S A R T Q S A CGACTGAGCACCGAGTGGGTGGGTGCGTATCTGGGGGGGG	1650 1700 1700 1700 1800 1800 1900
35 40	R A G V S S F G I S G T N A H V I CTGGAAAGGCACCCCCACTCAGCCTGCGGACAACGCGGTGATCGAGCG L E S A F F T Q F A D N A V I E R GGCACCGGAGTGGGTGCGTTGGGTGATTTCGGCCAGGACCCAGTCGGCTT A F E N V F L V I S A R T Q S A TGACTGAGCACGGGGTGATTGGGTGGTTACTGGGGGGGTGGCCCGGG L T E H E G R L R A Y L A A S P G STGGATATGGGGGGTGGTGGTGGTGATTCTGGGGGGGTGGCCCGGG L T E H E G R L R A Y L A A S P G STGGATATGGGGGGTGGTGGTGATGACACGGTGGGTGTT V C N H A V A S T L A N T F S V F CGAGCACCCTGCGGGGGGTGTGGTGACACGGGCACGGTG E H F A V L L G C D T V T G T A TGTCTGACCTTCGGGGGGGGTGTTCGCGGGGACAGGGTGGCGGGT V S D F P A V F V F P G Q G S Q R GTTGGCATGGGGGGACTGGCGGCGGGGGGGTGTCCCCGTTTCGGGGGGAT A J M G E E L A A A F P V F A R 1 CCATCAGCAGGTGTGGGACCTGCTGGATGTGCCGGATGTGAACG H Q I V W I L L D V P D L E V N AGACCGGTTAGGCCCAGCGGGCCCTGTTCGCAATGCAGGTGGCTCTGTTC E T G Y A Q P A L F A M Q V A L F	1650 1700 1700 1700 1800 1800 1900
35 40 45	R A G V S S F G I S G T N A H V I CTGGAAAGGCACCCCCACTCAGCCTSCGGACAACGCGGTGATCGAGCG L E S A F F T Q F A D N A V I E R GGCACGGGATGGGTGCGTTGGTGATTTCGGCCAGGACCCAGTCGGCTT A F E W V F L V I S A R T Q S A CGACTGAGCACGGGTGATGGGTGGGTGGTGATCTGGGGGGGG	1650 1700 1780 1880 1880 1900 1950 2000
35 40 45	R A G V S S F G I S G T N A H V I CTGGAAAGGGACCCCCCACTCAGCCTSCGGACAACGCGGTGATCGAGCG L E S A F F T Q F A D N A V I E R GGCACGGGATGGGTGCGTTGGTGATTTCGGCCAGGACCCAGTCGGCTT A F E W V F L V I S A R T Q S A CGACTGAGCACGGGCTGTTGGGTGGTTATCTGGGGGGGTTGGCCCGGG L T E H E G R L R A Y L A A S P G GTGGATATCCGGCTGTGGTGGTGGTGGTGATCTGGGGGGTTGGGTGTT V C M R A V A S T L A M T R S V F CGAGCACCCTGCGTGCTGGTGGTGATCACCGGCACCGCTT E H R A V L L G D D T V T G T A TGTCTGACTCTGGGGGGGTGTTCGTCTTCCCGGGAACGGGTGGCAGCGT W S D F R A V F V F P S Q G S Q R GTTGGCATGGGGGACTGGGGACCGGTTCCCCGTCTTCGCGGGAT A G M G E E L A A A F P V F A R I CCATCAGCAGGTTGGGACCTGCTGGTTCGCCGATCTGGAGGTGAACG H Q I V W I L L D V P D L E V N AGACCGGTTAGGCCCAGCGGGCCTGTTCGCAATGCAGGGGGTCTGTTC E T G Y A Q P A L F A M Q V A L F GGGCTGCTGGAATTGGGGTTACGACGGGCCCTTTCGGGCCATTC G L L E S W G V R P D A V I G H S GGTGGGTGAGCTTGCGGGTTGCGCGTTCGTTCGTGGGCCATTC G L L E S W G V R P D A V I G H S GGTGGGTGAGCTTGCGGGTTGCGCGTGTTCGTTTGGAGG V G E L A A A Y V S G V W S L E	1650 1700 1760 1760 1860 1860 1960 2000 2000 2100
35 40 45	R A G V S S F G I S G T N A H V I CTGGAAAGSSCACCCCCACTCAGCCTSCGGACAACGCGGTGATCGASCS L E S A F F T Q F A D N A V I E R GGCACCGGACTGGGTGCGTTGGTGATTTCGGCCAGGACCCAGTCGGCTT A F E W V F L V I S A R T Q S A TGACTGAGCACGACGGGCGGTTGGGTGGTTACTGGCGGGGCGTCGCCCGGG L T E H E G R L R A Y L A A S P G GTGGATATCCGGCCTGCTGGAGGACCGGACCGGTCGCCCGGG L T E H E G R L R A Y L A A S P G GTGGATATCCGGCCTGCTGGGAGATGACACGGTCGCCCGGT V C M R A V A S T L A N T B S V F CGAGCACCCTCCGGTGCTGCTGGGAGATGACACGGCACCGCTG E H F A V L L G C D T V T G T A TGTCTGACTCTCGGGCGGGTTCCCCGTCTTCGCGGGGAT A G M G E E L A A A F P V F A R 1 CCATCAGCAGGTGTGGGACCTGCTGGATGTGCCCGATCTGTTC E T G Y A Q P A L F A N Q V A L F GGGCTGCTGGAATCGTGGGGGTTGCCGGATGGGCCATTC G L L E S W G V R P D A V I G H S GGTGGGTGAGCTTGCGGTTGTCCGGGGGTTGGCCCATTCGGAGG V G E L A A A Y V S G V W S L E ATGCCTGCACTTTGGTGTTCGGGGGGTTGGCGGGGTTGTTGGCCC C A C L V S A B A R L M Q A L F	1650 1700 1750 1860 1860 1860 1960 2000 2000 2100 2150
35 40 45	R A G V S S F G I S G T N A H V I CTGGAAAGGGACCCCCCACTCAGCCTSCGGACAACGCGGTGATCGAGCG L E S A F F T Q F A D N A V I E R GGCACGGGATGGGTGCGTTGGTGATTTCGGCCAGGACCCASTCGGCT A F E W V F L V I S A R T Q S A CGACTGAGCACGGAGGACCCASTCGGTTGGTGATTCTGGGGGAGCCCASTCGGCT A F E W V F L V I S A R T Q S A CGACTGAGCACGAGGGCCTGTGGGTGATTCTGGGGGGGTCGCCCGG L T E H E G R L R A Y L A A S P G GTGGATATCCGGGTGCTGGTGATGAGCACGGTGGTGTT V C M R A V A S T L A M T F S V F CGAGCACCTTGCGTGCTGGTGAGAGCTGACCGGTACCGGTT E H F A V L L G D D T V T G T A TGTCTGACTCTCGGGCGGTTTCGTCTTCCCGGGAAAGGGGTCGCAGCGT V S D F R A V F V F P S Q G S Q R GTTGGCATGGTGGGAACTGGCCGGTTCCCCGTCTTCGCGCGGAT A G M G E E L A A A F P V F A R 1 CCATCAGCAGGTGTGGGACCTGCTGGTGCCCGATCTGGAGGTGAACG H Q I V W D L L D V P D L E V N AGACCGGTTACGCCCASCCGGCCCTGTTCGCAATGCAGGTGAACG E T G Y A Q P A L F A M Q V A L F GGGCTGCTGGAATTCGGGGTGTACGACGGGCCCATTC G L L E S W G V R P D A V I G H S GGTGGGTGACTTTGGGGTTGCGGGGGTCGTTGGAGG V G E L A A A Y V S G V W S L E ATGCCTGCACTTTGGTTTTGGTTTTGGGCTGTTGGAGG	1650 1700 1760 1760 1860 1860 1960 2000 2000 2100
35 40 45	R A G V S S F G I S G T N A H V I CTGGAAAGSSCACCCCCACTCAGCCTSCGGACAACGCGGTGATCGASCS L E S A F F T Q F A D N A V I E R GGCACCGGACTGGGTGCTTTGGTGATTTCGGCCAGGACCCAGTCSGCTT A F E W V F L V I S A R T Q S A TGACTGAGCACGAGGCCGGTTGGGTGCTTTCTGGCCAGGACCCAGTCSGCTT A F E W V F L V I S A R T Q S A TGACTGAGCACGAGGGCCGGTTGGGTGCTTACTGGCGGGGCTTGGCCCGG L T E H E G R L R A Y L A A S P G STGGATATCCGGGCGTTGGTGGTGGTGATGACACGGTCGCGGGTGTT V D M E A V A S T L A M T F S V F CGAGCACCCTGCGTGCTGGGAGATGACACGGTCGCGGGACCGGTG E H F A V L L G D D T V T G T A TGTCTGACTCTCGGGGGGAGTTCCCGGGGACAGGGGTTGGCGGGAT A G M G E E L A A A F P V F A R I CCATCAGCAGGTGTGGSACCTGCTCGGATGTGCCCGGTCATCGGCGGAT A G M G E E L A A A F P V F A R I CCATCAGCAGGTGTGGSACCTGCTCGGATGTGCCCGGTTGTTC E T G Y A Q P A L F A M Q V A L F GGGCTGCTGGAATCGTCGGGTTTCGGAATGCAGGGGTCGTCGTTC G L L E S W G V R P D A V I G H S GGTGGGTGAGCTTGCGGGTGTGCGGGGGTTGGCCGGTTTGGAGG V G E L A A A F P D A V I G H S GGTGGGTGAGCTTGCGGGTTTCGGGATGAGGGGTCGTTGGAGG V G E L A A A F P D A V I G H S GGTGGGTGAGCTTGCGGGTTTCGGGATGGAGGGGTTGTTGGAG V G E L A A A F P D A V I G H S GGTGGGTGAGCTTGCGGGTTTCGGGGGGTTGTTGGAGG V G E L A A A Y V S G V W S L E ATGCCTGCACTTTGGTGTTCCGGGGGGTTCGTCGGAGGGTTGTTGGCC C A D C L V S A B A R L M Q A L F GCGGGTGGGGGTGAGGGGGGGGTTGTTCGGAGGATGAGGGCTCTTGCCC C A C C L V S A B A R L M Q A L F GCGGGTGGGGGTGAGGGGGGGGGTCGTTGGGGGGGGGG	1650 1700 1760 1760 1860 1860 1960 2000 2050 2100 2150 2200

```
GRADE CREATURE CONTROL SE CONTROL SANCIA CARRA C
       TA TT 3201. 30600A.697.00033A.0A.036T0086TT0086C3A.30A.33T398C-2506
        TWVERVED TWEFFER
10
        TESAVETELSASES
       A R I T D G V A M I B G C B E I
      15
       A A I G A L A H L Y V N G V T
       RPALUSAAPATFVLI
       TGACATAGGCCTTGCAGCACCAGCGCTACTGGCTCGAGTCGGCACGCCCG
      PITARDEQRYWLESARE
20
      A A S D A S S P V L S S S I A L A
       03331000033300033973770400337700007337503375033700033
        BOBBBBTGTTCBTCBCCOMGBTGGCBCTGGCCGCGGCGAAGGGGGTCBAD BASA
25
      FAVEVABLALAADA.
      CATVERLDIASVEGFEG
      CCATGSCCGGACGACCGTACAGACCTGGGTCGACGAGCCGGCGGACGACG 3000
       H G R T T V Q T W V D E P A D C
30
      GCCGGCGCCGGTTCACCGTGCACACCCGCACCGGCGACGCCCCGTGGACG 3050
      G R R R F T V H T R T G D A P W C
      OTGCACGCCGAGGGGGTGCTGCGCCCCCATGGCACGGCCCTGCCCGATGO 3100
       L H A E G V L R P H G T A L P D A
      GGCCGACGCCGAGTGGCCCCCACCGGGCGCGGTGCCCGCGGACGGGCTGC 3150
35
       ADAEWPPPGAVPADGI
      CGGGTGTGTGGCGCGGGGGGGACCAGGTCTTCGCCGAGGCCGAGGTGGAC 3200
      PSVWRRGDQVFAEAEV:
      SGACCGGACGGTTTCGTGGTGCACCCCGACCTGCTCGACGCGGTCTTCTC 3250
      A P C C F V V H P D C L C A V F P
      DBDGGTDGGGGAAGGGGGGCAGGGGGGGGGATGGGGGAAGGTGADGG 3300
40
      AVGDGSRQPAGWRDL
      V H A S D A T V L R A C L T P F '
      SACGGAGCCATGGGATTCGCCGCCTTCGACGCCGCCGGCCTGCCGGTACT 3400
45
       DGAMGFAAFDGAGLPU
      CACCGUGGAGGCGGTGACGCTGCGGGAGGTGGCGTCACCGTCCGGCTCCG
        TAEAVTLREVASPSGS
      AGGAGTOUJACGGCCTGCACCGGTTGGAGTGGCTCGCCGAGGCG 2500 E E S D G L H R L E W L A V A E A
50
      GTCTACGACGGTGACCTGCCCGAGGGACATGTCCTGATCACCGCCGCCCCA 3550
       V Y D G D L P E G H V L I T A A H
      PDDFEDIPTRABTRAT
      GOGTOSTGACEGOCOTGCAACACCACCTCACCACCACCGAECACACCETE
55
      s v l m a l o H H L m m m D H m L
      ATCGTCCACACCACCACCGGCCGCCGCCCCCCTCACCGGCCTCAC 3700
       I V H T T T T D P A S A T V T S 1 T
      CODDAGOOOCAAAAAAAAAAAAAAAAAAAAAAAAAAA
        R T A Q M E H P H B I P L I E T
      ACCACCCCACACCCCCCTCCCCCTGGCCCAACTCGCCACCCTCGACCAC 3890
60
```

THE REPORT OF LARGE BASES OF HER r ta pa domino de la maño da Anexaño de maño a especial nome kobbe esta esta en la comencia de comencia de la comencia del comencia de la comencia del la comencia de la comencia de la CONTRATOR TOR CORRESPONDED TO SERVICE TO SERVICE TO SERVICE A SERVICE TO SERVICE TO SERVICE A SERVICE TO SERVICE A S a C M a c a m m b b s a m a a a 2300A27000330A270A22700001603A03T2633A000000A30AA0 4050 1.5 raari bogi beal mesk COMPRICATION OF THE PROPERTY OF A CONTROL OF THE CO L T T V L H P K A N A A W H L H ACCTCACCCAAAACCAACCCCTCACCCACTTCGTCCTACTCCAGCGCC 4250 20 3009003T00T0436A320003GA0AA3GAAA6TA0G00G003330AA630 4366 A A V L G S F G Q G N Y A A A U A COMMON TOGOD DE BONA TENERO DE CALONIO A CONTRA CONTRA CALONIA (A 4 D.C.) 25 T S I A W S M W B T T S T L T 3 2 Stogagsaccoccaggaccoartecgoogsscapttcottottat 4450 L D D A D R D R D R R D G F L F B CACGGACGACGAGGGCATGGGGATGCAT TEDEG 30

The AvrII-XhoI restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 13 (specific for methylmalonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below.

35. AGATOTGGGAGCTGGGGGAAGGGCTGGTGACGCTGGTGGGGGAGAGGAGCACC 50 I D A E A L D T D V R E S T GREGOROTOCTOGGREAGGTEGGTEGGTAGGACATCCCCGGGGACCGCGGC A A V L G H V G G E D I P A C A A GTTCAAGGACCTCGGCATCGACTCGCTCACCGCGGTCCAGCTGCGCAACG 150 40 FRDLGIDSLTAVQLES 000TCACCTCTOPPOPPANGCOPOTGTGCGGCTGTAACGCCADGGGGGTCTCCCAC 200 A L T E A T G V R L N A T A V F D TTOCCGACOCCGCTGCTOCCGGGAAGCTCGGGGAACGARCTGAC FFTFHVLAGKISCELTG 45 DADOU JOGOGOCCTOGTGCCCCSGACCGCGGCCACGGCCGGTGCGCACG 300 TRAFVVPRTAATAGAE D E P L A I V G M A C R L P G G V GOGTCACCOGAGGAGCTGTGGCACCTGTGGCATCCGGCACCGACGCCAT 400 50 ASPEELWHLVASGTDAI CACGGAGTTCCCGACGGACCGCGGCTGGGACGTCGACGCGATCTACGACC 450 TEFFTDRGWDVDAIYO OGGACCCCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500 PDPIAIGFTFVBGGFL 55 ADDGGDGDSADASGCTTOGACGDGSDGTTCTTCGGGATCAGCCCGDGCGA 550 T G A T G F D A A F F G I S P R E GUCCOTOGOGATGGACOGGCAGCAGCSGGTGCTCCTGGAGACGTCGTGGG 600

A L A M T F 2 2 F M L L E T S W
ASSOCITESAAASTITTESCATOACOCCOSACTOSACCCOCGOGGGGGAC 657
E A E E E S A S T T E E C D T F G S T
ASSOCITESATO COTOCCOTA COCTA CO F 1 0 Y F Y 3 1 1 1 A Y T V 10 A D S S S L D A L H L A G G S L P D COCCOSTOSCOC S B E C S L A L V G G V T V M A S P G G F V E F S R Q F G L A F S 15 3 F A K A . 3 A 3 A D 3 . 3 F A E GGGTGCCGGTGTGCTGATCGTCGAGAGGCTCTCCGAACGCCGAAACGCAAACG GAGVLIVERLSOA8911 STUACACOSTCOTUOCOSTOGTCCGTGGTTCGGGGGGTCAACCAGGATGGT 1100 20 S B T V I A V V B S S A V N Q I S OTOLANGGOGOTOTOGGGGGANGGGGGGGGGGGGAAGANGGAGGGGGGGAAT 1150 A S N G L S A F N G F S D E F V DOGGONGSODETBARQANGQQQGGGGCTCNDDCTGGGGGGGÄNDGTGGNDGCCCC (1290 25 FQALANASLTFABVIA V E A H S T G T R L G D P I B A Q 30GCTACTGCCCACCTACGGAGAGAGGGGGCCACCCCCCCCTGCTCCTGGG 1300 A V L A T Y G Q E R A T P L L L G 30 DIGGOTGRAGICONACATOGGOCACGOCAGGOCGGGTCCGGCGICGCCG 1350 S L K S N I G H A Q A A S G V A G I I K M V Q A L A H G E L P P T 35 L H A D E P S P H V D W T A G A V CGAAGTSOTGAGGTCGGGCCGGGCGTGGCCCGAGAGCGGGCCTAGGC 1500 E L L T S A R P W P E T D R P P GGGCGGGCGTGTCGTCGTTCGGAGTCAGCGGCACCAACGCCCACGTCATC 1650 RASVESEGVSGTUAEV GOA I AGEGGEACOCOCCECTCAGOCOSOGGAGGAGGGCGCAGCOTGTTGA - 1600 40 LESAPPAQPAEFAQPTE GACGCCGGTGGTCGCATGTGCTGCCGCTGGTGATATCGGCCAAGA 1650 T P V V A S D V L F L V I S A K DOCAGOOGGCCTGROCGAACACGAAGACTGGCTGDGCGCTACCTGGCG 1700 45 T Q F A L T B H B D E L F A Y L A TOGCCCGGGGCGGATATACGGGCCTGTGGCATCGACGCTGGCGGTGAC 1750 ASPGABIPAVASTLAVI ACGGTCGGTGTTCGAGCACCGCGCGTACTCCTTGGAGATGACACCGTCA 1800 RSVFEHRAVLLGGDTT COGGOACOGOGGTGACOGACOCCAGGATOGTGTTTGTCTTTCCCGGGGCAG 1850 50 TGTAVTDPRIVEPGQ GGGTGGCAGTGGCTGGGGATGGGCAGTGCACTGCGCGATTCGTCGGTGGT 1900 G W Q W L G M G S A L R D S S V V GTTCGCCGAGCGGATGGCCGAGTGTGCGCGGCGTTGCGCGAGTTCGTGS 1950 55 FAERMAECAAALREFV ACTGGGATCTGTTDACGGTTCTGGATGATCCGGCGGTGGTGGACCGGGTT 2000 D W D L F T V L D D P A V V D R V GATGTGGTCCAGCCCCTTGGGCGATGATGGTTTCCCTGGCCGGT D V V Q E A S W A M M V S L A A V GTGGCAGGCGGCGGTGTGCGGGCGGATGCGGTGATCGGCATTCGCAGG 2100 60

	W I A A G V A F F A V I G G H C [GI GAGATOR COLOAGOT TOT OF GEOGRAPO GEORGE COLOAC GACAN	215
-	- P. F. D. A. A. A. D. V. A. G. A. V. C. D. P. D. A. - P. C. GATOO MOOTO DO ANTAGO GATO OF COLUMN SAN	:
,	A B I I VII I LA B Z LA I A B X X X X X X X X X X X X X X X X X X	
	TOPACOGO TOPACOGO CO ACACO GO CO TOPACO TOPA	.30
10	- Presentation of the Architecture of the Arch	1.35
	A GRANDER A COMPANIE A	2400
15	ACGTOGAGOTGATOTGOGAGGAACTACTOGAGATGACTAGGGAGAGAGTAGT	2456
	TCGCAGACCCCCCTCCTGCCGTGCTGTCGACCGTGGACGGCAAAATTCCCCT	2500
	CGACAGCCGGCTGGACGGGAGTACTGGTACCGGAACCTGAACCGA	2550
20	D S P L D G E Y W Y R N L R E F TOGGTTTOCACCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	2600
	TTOGTTGAGGTCAGGGCCAGGCCGGTGTTGTTGCAGGGATGAAGGAAG	1651
25	F V E V S A S F V L L L A M D C C D TOTTOST LADGOTT COCACGOTT COTTOST CACGOTT COTTOST CACGOTT COTTOST CACGOTT COTTOST CACGOTT COTTOST CACGOTT COTTOST CACGOTT	1705
25	V V T V A T L R R D E 3 E A T F TGGTGAGGGGGGGAGGGGGTATGTGGAGGGGTGAGGGTGAACTGG	2750
	M L T A L A Q A Y V H G V T V D W GOOGEDATECTOGGEACCACCACACCGGGTACTGGACCTTCGGACCTA	2800
30	E À I L G T T T T R V L D L E T Y GEOTTECAACACCAGGGGTACTGGCTCGAGTCGGCACGCCGGGCGGCCGACT	2850
	A F Q H Q R Y W L E S A R P A A GESAGSSGGGGACCCCGTGGGGTGGGGTATCGCCCTCGGGGTGGGTG	2900
35	COGGGCCGGGTGTTCACGGGTTCCGTGCCGACCGGTGCGGACCGCGCGCG	2950
		3000
	CGGTCGAGCGGCTCGACATCGCCTCCGTGCCCGGCCGGCC	3050
40	T V E 8 L D I A S V P G R P G X G GGGGACGACGACGACGACGACGACGACGACGACGACGACG	3100
	R T T V Q T W V D E P A D D S E A COGGTTGACGTGACGTGACGTGACGTGACGTGACGTGACG	3150
45	COGAGGGGGTGCTGCCCCCATGGCACGGCCCTGCCCGATGCGGCCCGAC	3200
45	A E G V L R P H G T A L P D A A D GCCGASTGCCCCACCGGCCGCGCGCGCGCGCGCGCGCGCGCGCGCG	3250
	A E W P P E G A V P A C G L F S V GTGGCGCGGGGGGACCAGGTCTTCGCCGAGGCCGAGGTCGACGGACG	3300
50	W R R G E Q V F A E A E V D 3 F ACCOMPTEGRACICOSACCOSACCTSTTCGACCCCSACCTSTTCGACCCCSTCTTCTCCSCSSTC	3350
	D G F V V H P D L L D A V F S A V 330GACGGAAGCGGCCGGCCGGATGGCGGCGACTGACGGTGCACGC	3400
5 5	G D G S R Q P A G W R D L T V H A GTGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGG	3450
J J	CCATGGGATTCGCCGCCTTCGACGGCGGCCTGCCGGTACTCACCGGG	3500
	The state of the s	3857
60	GACGCCTGCACCGCTTGGASTGGCTCGCGGTCGCCGAGGCGACCCGC	3600

FILERUS WEAT ARATE ACTION AND BUT AND AND AND AND WELL AND Y ACTION OF COMMENCE OF CO i de la compania del compania de la compania del compania de la compania del la compania de la compania d tadandoa bahuububa baga badah baga badah baga batah baban Abbu (3.5%). a composa a a a compact a c 10 A C M E H P H A I P L I E T C E ikoApabanStrobadataSpabAA<mark>@to</mark>SabAap<mark>ttoG</mark>A@GA@Ga@iAob 3990 |H | T | P | L | F | L | A | Q | L | A | T | L | T | H | F | H TOURSON AND CONTRACTOR OF TOUR CONTRACTOR OF TOURS OF TOURS OF TOURS OF TOURS OF TOUR 15 Addaddaddaddadadaddaddaddaadaanaanaahadaddatdat 4000 CATCACOGGOGGOT TOGGGACOCTOGGCGGCATCOTCGCCCGCCGCCACC SA 4050 ITGGSGTLAGILARHI 20 M H P H T Y L L S A T P P P D A T PSTHLPCDV3DPHCLA CA OCOTOMOCOROR TO COCORRA B<mark>OCOTOMOCO COR TOTTOCROR</mark> COCOCORO HE CO T L T H I P Q P L T A I F H T A DEACOTOACOCTOACOCTOACO 4280 25 ATLEBGILHALT PDELI OSTCOTOCACCOCAAAGCCAMOGOCGCCTGGCACCTGCACCACCACCTCAC 4360 T V L H P K A N A A W H L H H L 30 Q N Q P L T H F V L Y S S A A A OTOGGOAGOCCOGGACAAGGAAACTACGCCGCCGCCAACGCCTTCCTC 4400 V L G S P G Q G N Y A A A N A F L GACGCCCTCGCCACCCACCGCCACACCCTCGGCCAACCCGCCACCTCCAT 4450 CALATHRHTLGQPATS C 35 OGCOTOGGGCATGTGGCACACCACCAGCACCCTCACCGGACAACTCGACG 4500 A W G M W H T T S T L T G Q L D ACSCOBACOGGGACOSCATODSCOGGGGGGGGTTTCCTCCCGGATCACGGAC 4550 40 GACGAGGGGATGGAT I E G

The *Nhe*II-*Xho*I restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 12 (specific for malonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below.

ASATOTSGCASCTOGCOSAAGCGCTGACGCTCGTCCGGGAGAGCACC E1

L A E A L L T L V R E S T

SCOSCOSTGCTCGGCCACGTGGGTGGCGAGGACATUCCCGCGACGGCGGC LCC

A A V L G H V S G E D I P A T A A

STTCAAGGACCTCGGCATCGACTCACCGGGGTCCAGCTGCGAACG LSC

F K D L G I D S L T A V Q L P W

CCCTCACCGAGGGGACCGGTGTGGGGCTGAACGCCACGGGGGTCTTCGAC LCC

A L T E A T S V E L N A T A V F E

TTCCCGACCCGGCACGTGCTCGCCGGGGAAGGACGAACTGACCGG LSC

F P T F H V L A G K L G D E L T G

CACCGGGGGGGGCCGTCCTGCCCGGGGGCCGATGGACGACG

```
THAR THAR STAR SH
           A F DO NA SIGA PODICE PROCEST PODICE PACE OF SIGNOCIA CONTRACTOR AND A CON
              23M200003h3000<del>h,</del>000bm3h000r000h000h000h000h000
           alpriate reveal of
 [1]
           A RECORDED BOOKS OF CORRECTION OF CONTRACT ACCORDED OF CONTRACT OF
           EAFECAGITEDETR
15
             Atronomico (nosabede<mark>ntancenadestade</mark>ntalentannesa 196
A 3 A 6 M e M e A 8 S M e M e T e A C
           DA GERA GERETT GERGEGERGEGGGGTGGGRERERA STETGGTGTGCRERER - 150
             T D G F G A T G S Q T S V L S 3
20
           R L S Y F Y G L E G P A V T Y D 1
           98679770<mark>07</mark>0670607667660660766766A66A660033R07666776667
           25
              S G E C S L A L V G G V T V M A
           300039<mark>30GAAGGGTTOGGGGGGGGGGGGGGGGGAGGTTOG</mark>GGGA
             3 R A K A F G A G A D G T S F A E
30
           SSGTGCCGGTGTGCTGATCGTCGAGAGGCTCTCCGACGCCGAACGCAACG 1050
             GAGVLIVERLSDAERU
           GTCACACCGTCCTGGCGGTCGTCCGTGGTTCGGCGGTCAACCAGGATGGT
           S H T V L A V V R G S A V M Q C
           SCOTOCAACGGGCTGTCGGCGCCCGAACGGGCCGTCGCAGGAGCGGGTGAT 1150
35
            A 3 N G L S A P N G P S Q E P V
           OCGBONGBOCOTGBOCANCSCOGGGCTCNCCCCGGGGGACGTGGNCGCGG
              R Q A L A N A G L T P A D V D A
           TOSAGGOCOMOSSOMOCOGGOACCAGGOTGGOCGMCCCCATCGAGGCACAS 1250
              GCGGTACTGGCCACCTACGGACAGGAGGGGCCACCCCCCTGCTGGG 1300
40
            AVLATYGQERATPLLLS
           CTOGOTGRAGTOCARCATOGGOCACGCCCAGGCCGCGTCGGCGTCGCCG 1350
             S L K S N I G H A Q A A S G V A
           SCAMONÍCANGAMBETBONGS DOCTOCGGONOGGENGOTGCCGCCGACG
45
           BII K M V Q A L R H G E L P P 1
           I H A D E P S P H V D W T A G A 7
           092240766764557666500093605766000346400540009300A0390
             E L L T S A P P W P E T D R P A
50
           STIGCOSCOTOTOTOTTOGGGGTGAGGGGGAGGAMGGCCAAGGTCKTC 1550
          B A A V S S F G V S G T N A H V
           CTGGAGGCCGGACCGGTAACGGAGACGCCCGCGGCATCGCCTTCCGGTGA 1600
            leastrampaas ps si
           COTTOCCOTGOTGGTGTGGGCACGCTCACCGGAAGCGCTCGACGAGCAGA 1680
55
             L F L L T S A R S P E A L D E C
          TOOGGOGACTGOGGGCTACCTGGACACCACCGGGACGTCGACCGGGTG 1700
           AVAQTLARRTHFAHRAV
60
           GOTGOTOGGTGACACOGTCATCACCACACCCCCGGGGACCGGCCCGACG 1800
```

5 atyada odno ot odanat abadaat ot adadatiko daakko dattada (1985 World to the botton and a contract to padagobagoobtahtoopaataoaggeerotatrabagotabagaa.2000 A. Q. F. A. D. F. A. M. Q. M. A. D. F. G. D. D. B. argaaarana gaachalaadagaraardagala rhoolin allamuaalin 2053 10 TODOTAMOTOTOOSSSONGTOSTOSTTOSASSATSATTOOACTT 2105 A A A Y Y S S Y W S D E D A D SGTGT nggaggaggagataanan ganggat at gaadaa ku baggata (159 15 1 T S A A A F L M Q A L F A C C T ATSATICECTOTECCESTOTEAMEMEMEMEMESCECEESCOSTECTESETEM 2200 M V A V P V S B D B A B A V L G B G V B I A A V N G P S S V V L S 20 373A70A3390G0CGCCCCCCAGGCCGCGAAGGSCCTGGGAAA37GGAG3 2366 G D E A A V L Q A A E G L G F W T RLATSHAFHSAFMEFMI 99%99%3TT00999096T09009AM6600T0%00T%000A031030AG3 2400 25 TOTOCATGGCCGTTGGTGATCAGGTGACCACCGCTGAGTACTGGGTGCGG 2450 V S M A D G C Q V T T A E Y W V F CAGGTCCGGGACACGGTCCGGTTCGGCGAGCAGGTGGCCTCGTACGAGGA 2500 Q V R D T V R F G E Q V A S Y E D OSCOGTETTOSTOGRACTEGGTGCCGACCGGTCACTEGCCDECCTGGTCG 2550 30 AVFVELGADRSLARLV ACGSTGTCGCGATGCTGCACGGCGACCACGAAATCCAGGCCGCGATCGGC 2600 D G V A M L H G D H E I Q A A I G GUCCTGGCCDACCTGTATGTCAACGGCGTCACGGTCGACTGGCCCGCCT 2650 A D A H L Y V W G V T V D W F A L 35 CONGGGGGATTGCTOCGGGAMONOGGGTGCTGGACCTTCCGACATACGCCT lgyagaravlbletya TCCAGCACCAGGGGTACTGGGTCGAGTCGGCACGCCGGGCGCATCCGAC 2750 FQHQBYWLESAFFAASC DOGGETERROCCETTACTOROTATOGGETATOGGETERCOGGETER 40 AGHPVLGSGIALAGSPG R V F T G S V P T G A D R A V F V A E L A L A A A D A V E C A T V 45 GASCGGCTCGACATCGCCTCCGTGCCCGGCCGGCCGGGCCATGGCCGGAC 2950 ERLDIASVPGRPSHSR¶ GACCETAGAGACCTGGCCGACGAGGGGGGGACH CTGCCGGGCGGGT 3000 TVQTWVPEPADDGRRR 50 TONCOGTGCACACCCGCACCGGGGGGGGGCGCGTGGACGCTGCACGCGAG 3050 FTVHTRIGDAPWILEAE SGGGTGGTGGGGGGGGATGGGAGGGGCCTGCCGGATGGGGGGGAGGGGGA 3 V L R P H G T A L P I A A I A É STEROCOCCACCESSOCGSTGCCGCGGGGACGGGCTSCCGGGTSTGTGGC 3150 55 were road veabore row 300000300ACCAGGTCTTC3CCGAGGCCGAGGTGGACGGACCGGACGGT 3200 P A D D Q V F A B A B V D G P D G Tredrestronessonancerastrognogogoterrorscogogoressoga 3250 F V V H F D L L D A V F S A U G D 60 SGGAAGCCGCCAGCCGGCCGCATGGCGCGACCTGACGGTGCACGCGTCGG 3300

WO 00/20601

GRAND WIND TO THE AND NA ARANG BRANG BUBGGGGGGGGGGGGGGAGGGAAGGGAAGGGAAGG A T V D B A C B T B B D B A M etalot ninniberberake as hisopopolotoo asta belabe hoolaasa ee ee ee s Oo ee aa aa ee oo oo aa aa boo boo oo aa aa aa 33.44.332734633453,353<mark>37777</mark> T T L A A T A S P S A S B B S acordonados artes estos de producios de la composição de 1900 d o la fila wil a vala vito 10 COASSACATACECACESSESESESEACCASCESCACEGOSTECTEACES 3600 E C : F T F A H T R A T R V L : 15 NEBPBBIRLIETDBBB 00000T0000Cpgg000AACp<mark>eG00ACCC</mark>TGGACACCCCCACCTCC3CC 20 TO FOL POLIA QUE A TOLD HOP HOLD FOR THE CONTRACTOR SACRETOR SACRE OA COCCA COCA COTOCA COCCCCTCAACCCCCAAA CA COCCATOA TOA TOA COA 25 'PPTTTPLNPEHALLI COGGCGCTCCGGCACCCTCGCCGGCATCCTCGCCCGCCACCTGAACCAC 3950 T G G S G T L A G I L A R H L M H PHTYLLSRTPPPDATPG 30 DACCCACCTCCCCTGCGACCTCGGCGACCCCACCAACTCGCCACCACCC 4050 THLPSSVGDPHQLATT TCACCCACATCCCCCAACCCCTCACCGCCATCTTCCACACCGCCGCCGCACC 4100 LTHIPQELTAIFHTAAT OTOGACGACGGCATOCTOCACGCCCTCACCCCCGACGGCCTCACCACCGC 4150 35 LEDGILBALTPERLITY COTOCA CODORAMOCOMA CASCAGA COTAGA A COTAGA CACATOR COCARA - 4200 L H P E A N A A W H L H H L T ADDALODDSTENDESNOTTSSTOCTCTACTCCASDGCCGCCGCCGTUCTC 4250 D D P L T H F V L Y S S A A A C C 40 930AGCCCGGACAAGGAAACTACGCCGCCGCCAACGCCTTCCCCCAACGC 4300 G S P G Q S N Y A A A N A F L D A CCTCGCCACCCACCGCCACACCCTCGGCCAACCCGCCACCTCCATCGCCT 4350 LATHRHILGQPATSIA GGGGGATGTGGGACACCACCAGCACCCTCACCGGACAACTCGACGACGCC 4400 45 W G M W H T T S T L T G Q L D D A GACCGGGACCGCATCCGCGGCGGCGGTTTCCTCCCGATCACGGACGACGA 4450 D R D R I R R G G F L P I T D D E GGGCATGGGGATGCAT

50

The *Nhe*II-*Nho*I restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 13 (specific for methylmalonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below.

AGATOTOGOASOTOGOGGARGOGOTGATGGTGATGGGGAGAGCAGC 50
Q L A E A L L T L V R E S T
GCCGCCGTGCTCGGCCACGTGGGTGGGGAGGACATCCCCGGGAGGGCGCG100

	A A V 1 3 B V 7 3 B E 1 1 1 A A A A A A A A A A A A A A A A	
5	 Continuo nuo pera e dum otronos trigala de tra ograpo et degale. Al continuo en la como esta en la c	
	TIPO PARO POSA PER NO LO A GOMENTO SE POR EL TORA	250
	- ANDERSONS TO THE STATE OF THE	300
10	DEAEVVPETAATAGAE	
10	ACGMGCCGCTGCCGATCGTGGGMATGCCCTGCCGGCTGCCCGGCGCGGGGG	950
	SCSTCACCISA SGAGOTGTGGCACCTCGTGGCATCCGGCACGGCGCCAT A S P E E L W B L V A S G T C A I	400
15	CACSSAGTTUDEGACGGACCGCGGGCGGGGGGGATCTACGACC	: 50
. `	CASA DO CARRAGA DA TRABA A CONTROCTO CA SUA CARRAGA PARA CONTROCTO	500
	AppagagganeAggarradghaggggarrattighrahgagggah	550
20	TO BEACT BE DEALER FOR BUSINESS	600
	A L A M D P Q Q R V L L S T S W AGGISTTOGARAGOCOGGCATOROCOGGATTOGARCOCOGGCATOROCOGGATTOGACCOGGCGGCGASCGAC	22.5
	E A F E S A G I T P D S T R G 3 D	200
25	ACCOMINGTOTOGTOGGCGCCTTCTCCTACGGTTACGGCACCGGTGCGGA T G V F V G A F S Y G Y G T G A D	7 Gre
	CACCGACGGCTTCGGCGCGCGCCTCGCAGACCAGTGTGCTCTCCGGCC T D G F G A T G S O T S V L S G	750
	GGCTGTCGTACTTCTACGGTCTGGAGGGTCCGGCGGTCACGGTCGACACG	800
30	GCGTGTTCGTCGTCGCTGGTGGCGCTGCACCAGGCCGGGCAGTCGCTGCG	850
	CTCCGGCGAATGCTCGCCCCCGGCGCGCGCGCGCACGGTGATGGCGT	900
	S G E C S L A L V G G V T V M A COTCCCGGGGGGCCTCGCGCCCGGGAGC	950
35	S P G G F V E F S R Q R G L A P D	1000
	GSCCGGGCGAAGGCGTTCGGCGCGGGGGGGCACGACCTTCGCCGA G E A K A F G A S A D G T S F A B	1000
•	GREAT SCORET OF THE CONTROL OF THE C	1050
40	STOR CACOSTOCTEGOCGTCGTCGCTCGTCGCTGGTTCAACCACATGST	1100
		1150
	A S N G L S A P N G P S Q E R V I COGSCAGGCCCTGGCCAACGCCGGGGTGACCCCGGGGGACGTGGACGCCG	1200
45	R Q A L A N A G L T P A D V D A	1250
	TOGAGGCCACGGCACCAGGCTGGGCGACCCCATCGAGGCACAG	1250
	GOGGTACTGGCCACCTACGGACAGGAGGGCGCCACCCCCCTGCTGCTGGG	1300
50	A V L A T Y G Q E R A T P L L L G CTCSCTGAAGTCCAACATCGGCCACGCCCAGGCCGGGTCCGGCGTCGGCG	1350
	S L K S N I G H A Q A A S G V A GCATCATCAAGATGGTGCAGGCCTTCCGGCAGGGGAGCTGCCGCCGACG	1400
	G I I K M V Q A L F H G E L F F T T TO THE CACGOOGLOGGEST TO THE CONTROL OF THE C	1450
55	LABEPSPHVDWTAGAV	
	E L L T S A P P W P E T F P R	1500
	STECCECCECTCCTCSTTCSSESTSAGGGCACCAACGCCCACSTCATC	1350
60	OTGGAGGGGGAGGGGTAAGGGAGAGGGGGGGGGATGGGGTTGAGGTGA	1 - 20

I PUTETAA. 10 AACTISTOTTOGTOTACTOGGGGGAGGGGAGCAAGAATICCGGGGGATSGGG 1550 ELVEVYS GQGTQBBPAMG GAGCAGCTAGCCGATTCGTCGGTGTTCGCCGAGTGGTCGAGTGIL900 R Q L A D S S V V F A E R M A E PGOGGOGGETT SCOCSAGTT CGTGGACT GGGAT CTGTT CACGGTT CTGG 1953 A A A L R E F V D W D L F T V L 1.5 ACPATOCOGOGOTOS, JACOGOS: JATGTCOTOCAGOCOGOT.COCTOU 2000 D D A V J D R V D V U 1 P A 3 N 326ATGATGGTTT000T3G00GCGGTGTGGCAGG03G0GGGTGTGTGDG3G00-2050 A M M V S L A A V W Q A A G V A B 20 200617300677370207200327600<mark>0000000277076200</mark>075200230 A G A V E L R D A A R I V T : 25 D A I A R G L A G R G A M A S V A CTGCCCCCCAGATGTCGMGCTGGTCGACGGGGCCTGGATCGCCGCCC 2250 L P A Q D V E L V D G A W I A A ACAACGGGCCCGCCTCCACCGTGATCGCGGGCACCCGGGAAGCGGTCGAC H N G P A S T V I A G T P E A 7 D 30 CATGTCCTCACCGCTCATGAGGCACAAGGGGTGCGGGTGCGGCGGATCAC H V L T A H E A Q G V R V R R I I CSTOGACTATGCCTCGCACACCCCGCACGTCGAGCTGATCCGCGAGGAAC V D Y A S H T P H V E L I R D E 35 J L D I M S D S S Q M P L 7 P W omandakodattagkodadakonggutogkokaddosottagkoggagaketk 2500 L S T V D G T W V D S P L D G E T OTGGTN003GAA deTgogTgAM@GGGTCGGTTTCCX0CCCGGCGTCAGCC W Y R N L R E P V G F E P A V S 40 AGTTGCAGGCCAAGGGGGACACGGTGTTCGTCGAGGTCAGCGCCAGCCCG LIAGGOTVEVEVSAS GTGTTGTTGCAGGCGATGGACGA<mark>CGATGTCG</mark>TCACGGTTGCCACGCTGCG 2650 T L L Q A M D D D V V T V A T L R OGTGAOGAOGGOGAOGGOACCGGATGCTCACCGCCCTGGCACAGGCCT 2700 45 R D D G D A T R M L T A L A Q A ATGTCCACGGCGTCGACTGGCCCGCCATCCTCGGCACCACCACACA 2750 Y V H G V T V D W P A 1 L G T T T ACCCGGGTACTGGACCTTCCGACCTACGCCTTCCAACACCAGCGGTACTG 2800 TRVLDLPTYAFQEQRYX 50 GUTUGRETURGUAGGOOGGCGGCATCCGACGGGGGCACCCCGTGCTGS 2850 L E S A R P A A S D A G H P V L GCTCCGGTATCGCCCTCGCCGGGTCGCCGGGCCGGGTGTTCACGGGTTCC 3 S G I A L A G S P G R V P T G S 379003400367906340066606<mark>676776</mark>67060064667660567580 1356 55 V P T G A D R A V P V A B L A L A DBCCGCGGACGCGSTCGACTGCGCCACGGTCGAGCGGCTCGACATCGCCT 3000 A A C A V C C A T V S P L C I A S V P G R P G H G R T T V O T W V 60 GACGAGCOGGCGGHCGACGGCGGCGCGGGTTCHCCGTGCHCACCCGGCAC 2100

LEE FALL FOR BEFEEL WEST BUT STORES TO STORE STO a ah nadh sheba bayba ah sababar ah sababa sababa abeedeen ah basa b Walio Alio aha oo aha oo aha oo ahaa ahaa waxaa walio ahaa baaba ah ahaa ahaa ah December of the strain of the control of the contro ABABUSGPDSFUVBFS WOOTBGARDSCOTTOTGGGGGGTGGGGGAAGGGGAAGCGGCAGCGGGCC 33E0 10 SWF SUTWHASDATUDEA UTO TOTTA DO DE SOU DA DO DA A OGRAFICA TO DO DO DE TO DA ON 1948 ; 15 abannanahmin pagmwemekee<mark>gaagkadoogm</mark>gwegomacaaakaama (380) G A C L B W L T A B A W D L SAUTOACOSTOCESOTOCEAGGACTOGGACCGGCTGCACCSCTTEGAETG 355) A S P S S E E S D G L H F L E W GOT OG COSTOS COSAGOO GOTOTA COA COSTO A COTO COSA SESSA CATE (3600 20 LAVABAVY B G E L P E 9 H TOOT SATISATION OF TO GOOD ACCORDANCE THE CARROLS OF THE CARROLS OF THE CONTROL O 25 A H T R A T R V L T A L Q H H L T T T D H T L I V H T T D P A G CCACCGTCACCGGGCTCACCGGCACCGCCCAGRACGRACACCCCCACCGC 3800 ATVTSLTRTAQNEHPHR ATCCCCCTCATCCAAACCCACCCCCACACCCCCCCTCCCCCTGSCCCA 3850 30 IRLIETOHPHTPLPLAQ L A T L D H P H L R L T H H T L 35 H H P H L T P L H T T T P P T T T COSOT CAA COODGAMCA OGCCATCATCATCACCGGCGGCGCTCCGGCACCCT FLMPEHAIIITGGSSTL CONSTRUCTION OF THE PROPERTY O 40 GTOSSCSACCCCCACCAMOTOGOCACCACCCTCACCCACATCCCCCAACC V G D P H Q L A T T L T H I F Q P corologaciano in tollogacia con con contra de la contra del contra de la contra del la contra de la contra de la contra del la contra del la contra de la contra de la contra del la contra del la contra de la contra de la contra del la LTAIFHTAATLDOGIL 45 AGGCCCTCACCCCGACCGCCTCACCACCGTCCTCCACCCCAAAGCCAAC 4250 ograciataciacagogogogogogogotogoagaacaagaaa 4350 50 VLYSSAAAVLGSPSQG U Y A A A N A F L D A L A T H A H ACCCTCGGCCAACCCGCCACCTCCATCGCCTGGGCATGTGGCACACCAC 4450 55 TLBQPAISIAWGMWHT' CAGCACCOTOÁCOGGACAACTOGACGACGCCGACCGGGACCGCATCCGCC R G G F L F I T D D E G 60

15

20

25

30

Phage KC515 DNA was prepared using the procedure described in Genetic Manipulation of *Streptomyces*. A Laboratory Manual, edited by D. Hopwood *et al.* A phage suspension prepared from 10 plates (100 mm) of continent plaques of KC515 on *S. lividans* TK24 generally gave about 3 µg of phage DNA. The DNA was ligated to circularize at the cos site, subsequently digested with restriction enzymes *BamHI* and *PstI*, and dephosphorylated with SAP.

Each module 8 cassette described above was excised with restriction enzymes *BgI*II and *Nsi*I and ligated into the compatible *Bam*HI and *Psi*I sites of KC515 phage DNA prepared as described above. The ligation mixture containing KC515 and various cassettes was transfected into protoplasts of *Streptomyces lividans* TK24 using the procedure described in Genetic Manipulation of *Streptomyces*. A Laboratory Manual edited by D. Hopwood *et al.* and overlaid with TK24 spores. After 16-24 hr. the plaques were restreaked on plates overlaid with TK24 spores. Single plaques were picked and resuspended in 200 µL of nutrient broth. Phage DNA was prepared by the boiling method (Hopwood *et al.*, *supra*). The PCR with primers spanning the left and right boundaries of the recombinant phage was used to verify the correct phage had been isolated. In most cases, at least 80% of the plaques contained the expected insert. To confirm the presence of the resistance marker (thiostrepton), a spot test is used, as described in Lomovskaya *et al.* (1997), in which a plate with spots of phage is overlaid with mixture of spores of TK24 and phiC31 TK24 lysogen. After overnight incubation, the plate is overlaid with antibiotic in soft agar. A working stock is made of all phage containing desired constructs.

Apr 1966, incorporated herein by reference) mycelia were infected with the recombinant phage by mixing the spores and phage (1 x 10⁸ of each), and incubating on R2YE agar (Genetic Manipulation of *Streptomyces*, A Laboratory Manual, edited by D. Hopwood *et al.*) at 30°C for 10 days. Recombinant clones were selected and plated on minimal medium containing thiostrepton (50 μg/ml) to select for the thiostrepton resistance-conferring gene. Primary thiostrepton resistant clones were isolated and purified through a second round of single colony isolation, as necessary. To obtain thiostrepton-sensitive revertants that underwent a second recombination event to evict the phage genome, primary recombinants were propagated in liquid media for two to three days in the absence of thiostrepton and then spread on agar medium without thiostrepton to obtain spores. Spores were plated to obtain about 50 colonies per plate, and thiostrepton sensitive colonies were identified by

WO 00/20601

10

15

20

25

30

35

replica plating onto thiostrepton containing agar medium. The PCR was used to determine which of the thiostrepton sensitive colonies reverted to the wild type (reversal of the initial integration event), and which contain the desired AT swap at module 8 in the ATCC 14891-derived cells. The PCR primers used amplified either the KS AT junction or the AT:DH junction of the wild-type and the desired recombinant strains. Fermentation of the recombinant strains, followed by isolation of the metabolites and analysis by LCMS, and NMR is used to characterize the novel polyketide compounds.

93

Example 2

Replacement of Methoxyl with Hydrogen or Methyl at C-13 of FK-506

The present invention also provides the 13-desmethoxy derivatives of FK-506 and the novel PKS enzymes that produce them. A variety of *Streptomyces* strains that produce FK-506 are known in the art, including *S. tsukubaensis* No. 9993 (FERM BP-927), described in U.S. Patent No. 5,624,852, incorporated herein by reference: *S. hygroscopicus* subsp. *yakushimaensis* No. 7238, described in U.S. patent No. 4,894.366, incorporated herein by reference; *S.* sp. MA6858 (ATCC 55098), described in U.S. Patent Nos. 5,116,756, incorporated herein by reference; and *S.* sp. MA 6548, described in Motamedi *et al.*, 1998, "The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK-506," *Eur. J. Biochem. 256*: 528-534, and Motamedi *et al.*, 1997, "Structural organization of a multifunctional polyketide synthase involved in the biosynthesis of the macrolide immunosuppressant FK-506." *Eur. J. Biochem. 244*: 74-80, each of which is incorporated herein by reference.

The complete sequence of the FK-506 gene cluster from *Streptomyces* sp. MA6548 is known, and the sequences of the corresponding gene clusters from other FK-506-producing organisms is highly homologous thereto. The novel FK-506 recombinant gene clusters of the present invention differ from the naturally occurring gene clusters in that the AT domain of module 8 of the naturally occurring PKSs is replaced by an AT domain specific for malonyl CoA or methylmalonyl CoA. These AT domain replacements are made at the DNA level, following the methodology described in Example 1.

The naturally occurring module 8 sequence for the MA6548 strain is shown below, followed by the illustrative hybrid module 8 sequences for the MA6548 strains.

```
SUMTSOSSOTSTACCAGGOSSSACGGGGACGGGAAGT COUGTOGTGGTG 50

M.R. D. Y. E. A. A. R. R. T. G. S. P. V. V. V.

GCGGCCGGGCTUGACGACGCGCGCGGACGTGCCGCTGCTGCGGGGGCTGCG 100

A. A. A. D. D. A. F. D. V. F. D. L. P. G. D. R.
```

```
ovala a a vale e a socialis
                                  ast so mak mak pake pa pa pake a tan be b T
               CON SOLA NA SANTAS DAD CATON DO COLATON BORGO DA ELA RACÁN (CENTRO). NO LOS COLATON DE LOS CALOS DEL LOS CALOS DE LOS CALOS DE LOS CALOS DE LOS CALOS DE LOS CALOS DEL LOS CALOS DEL LOS CALOS DE LOS CALOS DE LOS CALOS DEL LOS C
                0003303A33A3ACSACSACSAAS<mark>3AACT</mark>OGSAATCSACTOSCTCACGSGS
                   FAITOPPEBLOICELA
               10
               LA V P D P P T P R A L A A P L S
               DGA GGA GOT GGGGGGTA CCCGGGGGGGGGGGGGGGGGGGGGGGAACCGGGGGAA
                  D E L A G T R A F V A A R T A A
               DIGGGGGGGGGAACGAAACGGCTTGGGGAACGTGGGGAATGGGGAATGGGGT
15
                  AAAHLEPLAI73MA
               TTGCCGGGGGGGGGTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC
                 LPGGVASPQELWRLUAS
                20
                  G T D A I T E F P A D R G W D V
               A 330GOTOTA OGA QOOQGA GOOGA OQOGA TOGGOAA GA OOTIYOGI GOGG - 650
                 25
               GATCAGCCGGCGCGAGGCCCTGGCCATGGACCCGCAGCAACGGGTGCTCC 750
                 ISPREALAMOPQQFUS
               DOGNONDE CONCOUNDING BARBOOTT CORRESPONDE CONCOUNTE DE CO
                 ETSWEAFESAGITFOA
               BOGCGGGGCAGCGACACGGGGGTGTTCATGGGCGCGTTCTCCTACGGGTA 850
30
                ARGSDTGVFIGAFSYGY
              CGGCACGGGTGCGGATACCAACGGCTTCGGCGGGCGACAGGGTCGCAGACCA 900
                 STGADTNGFGATGSQT
               GOGTGOTCTCCGGCCGCCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950
              SVLSGRLSYFYGLEGFS
35
               STCACGGTCGACACCGCCTGCTCGTCGTCACTGGTCGCCCTGCACCAGGC 1000
               V T V D T A C S S S L V A L H Q A
              AGGGCAGTCCCTGCGCTCGGSCGAATGCTCGCTCGCCCTGGTCGGCCGTG 1050
                 TCACCCTGATGCCCTCCCTCCCCCCCCACTTCCTCGACTTCTCCCCCCAAAGCCC 1100
40
               ITVMASPSGEVEESRQR
               G L A F C G A A K A F G A G A D G
              TACGAGOTTOGOOGAGGGOOGGTGCCCTGGTGGTCGAGCGGCTCTCCC 1200
                 TSFAEGASALVVERLS
45
              DARREGHTVLAL/ROSA
              SCTAACTCCGACGGGGGTCGAACGGTCTCTCGGGGGCGAACGGCCCCTC 1300
               ANSDGASIGLSAPNGPS
               CCAGGAACGOGTCATCCACCAGGOCCTCGOGAACGOGAAACTCACCCCCG 1350
50
               QERVIHQALANAKLTP
              COGATGTOGACGOGGTOGAGGOGTACGGCACCGGCCACCGCCTCGGCGAC 1400
              A D V D A V E A H G T G T R L G D
               CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450
               PIBAQALLATYS Q D R A D
               GCCCCTGCTGCTCGGCTCGCTGAAGTCGAACATCGGGCACGCCCAGGCCG 1500
55
                  PLLLGSLKSNIGHAQA
                OTCAGGGGTCGCCGGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550
              A S G V A G I I K M V Q A I B H G
              GAACTGCCGCCGACACTGCACCCGGACGAGCCGTCGCCGCACGTCGACTG 1600
60
                E L P F T L H A D E P 3 P H V D W
```

5 AARGOODAAN ON OOTSAGGAAGGGGTAAAA GGGGACGGGTGA INSC | W. A. H. C. C. C. E. A. G. B. W. R. T. | W. E. A G A T E A G F V B V G F J B A 10 PRUPABERS AFB SCIES PT GBT 6TG 6BG 6BGT TGG GG BAGAAAT GGA GGAAAAT GYYG GGAGAT V S A B S F E A L D E Q I S E L agagoramoregaoaseseseseserasas as geseses contesce R A Y 1 D T 3 P G V D F A A V A 15 AGAGACTESCOCGGCGTAGGCACTTCACCCACGGGGCGTACTGCTCGGG 2000 Q T L A R R T " F T R R R W L L G GAPT COSTOTT" TOTACTOCGGTCAGGGCACCCAGCATCCCGCGATGGGCGAGCAACTCG 2100 20 T Y S G Q G T Q H P A M G E esseessettedeegteteseeghtseettsekeskassasteesekesk A A A F F V F A D A W B I A I B P DTOGROGROOOGROOCGGROGROORGGROEGROOFSCKORGGRONDOGTOTT 2200 osocoaccagoogocttchooscoctcotgaggcotosga/arcacg 25 A H Q A A F T A L L R 3 W L I T cochogoogranceecontogoresaesaancaecoocesaaces P H A V I G H S L G B I T A A Y A GODGGGATCCTGTCGCTCGACGACGCCTGCACCCTGATCACCACGCGTGC 2350 A 3 I L S L C D A C T L I T T R A COSCOTOATGCACACGCTTCCGCCGCGCGCGCATGGTCACCGTGCTGA 2400 30 $\texttt{R} \quad \texttt{L} \quad \texttt{M} \quad \texttt{H} \quad \texttt{T} \quad \texttt{L} \quad \texttt{F} \quad \texttt{P} \quad \texttt{P} \quad \texttt{G} \quad \texttt{A} \quad \texttt{M} \quad \texttt{V} \quad \texttt{T} \quad \texttt{V} \quad \texttt{L}$ CCASCGAGGAGGAGGCCCTCAGGCGCTGCGGCCGGGCGTGGAGATCGCC 2450 T S E E E A R Q A L R P G V E I A 35 SUSSMOTTEGECCCGCACTCCSTCGTGCTCTCGSGCGACGACGACGCCGT 2500 A V F G P H S V V L S G D E E A V GETEGACGTEGEACAGEGETEGGEATECACGACGGTETGCCGGGGGGGG 2550 L D V A Q E L S I H H F L F A P #38333800#CT006032#6#T33##80009T638800003#33T30T0300 40 B A G B S A B M B P V A A B L L A ACCACTOGOGAGOTOCOTTACGACOGGCCCCACACCGCCATCCCGAACGA 2650 TRELRYDRPHTAIPND roccaccacceceastactesecceaecaccectescaaccccstcotet 2700 TCCACGCCAACCCAGCGTACCCCGACGCCGTCTTCGTCGAGATCGGC 2750 45 F H A H T Q R Y P D A V F V E I G CCCGGCCAGGACCTCTCACCGCTGGTCGACGGCATCGCCCTGCAGAACGG 2800 PSQDLSELVDGIALONG CACGGCGACGAGGTGCACGGGCTGCACACGGGGCTCGCCCGGCCTCTTCA 2850 50 TADEVERLETALARDE CACGOGGCGCCACGOTCGACTGGTCCCGCATCCTCGGGGGTGCTTCGCGG 2900 CACGACCCTGACGTCCCCTCGTACGCGTTCCAGCGGCGTCCCTACTGGAT 2950 H D P D V P S Y A P Q P R F Y W I CGASTOGGOTOCCCGGGCACTGGGGCCACCCCGTCCTGGGCA 3000 E S A P P A T A D S G H P V L G 55 COGRAGMOGOCOMOGOCAGOMOGOCGGGCCGGGMAMAMACGGGTCCCGTG 3050 COCGCCGGTGCGGACCGCGCGTTTCATCGCCGAACTGGCCGCCG100 60 PAGABFAVETABLADAA

WO 00/20601 PCT/US99/22886

96

```
TA DIA TIBLO A TIMB LI LI MIT S
BEGGGGGGGGGGGGGGGAGGGGCACTGGGGTGGAT SACC
                           official property and a second control of the second control of th
                ngA digon do amagA digong GA de de de de asmiron do godo niverbigo a la signi
I Din Alian Militan Lingua Alian Billian Militan Billian Billian Billian Billian Billian Billian Billian Billia
                10
                20090<mark>038A03990T002GGGGGGGTGGGGA</mark>ESC2TGBWGGABAXXXXXX
                 PADBLEGANEFAL,
               037430034461034046000134660770313304640000000000003451
                   E A E V D S P D G F V A E F C
               DOGNOGOGOTOTTOTOGGOGGTOGGGGACHAN AAAAAAGCCGACAA 3890
15
                Logia to the solar volume of the control of the con
                W R D L A V H A S D A T V L F A C
                COTORCOCCOCCARA <mark>TEGTET GETGA GOT COCCOTT</mark> COR 1901 GET GA 300 G
20
                    L T R R D S G V V E L A A F 1
               DOGGAATGCCGGTGCTCACCGCGGAGTCGGTGACGCTGGGGGGAGGCCGCGGGGGA
               AGMEVITAESVOLSEV
               TOBGOAGGOGGATOGGACGACTOGGAOGGTOTGCTTOBSCTTGACIGSCTT
                    RCCGGTGGGGGAGGGGGACTAG<mark>GAGGGTG</mark>GGGAAGGAGGTGGGGAAGGAGGG
25
                   F V A E A H Y D G A D E 1 F E G
              30
                PHNTPTRTHTQTTRVLS
               OGCCCTCCAACACCACCTCATCACCACCAACCACACCCTCATCGTCCACA 3900
                 ALQHHLITTNETLIUH
               T T T D P P G A A V T G L T B T A
               CAMAACGAACROCCGGCGGCATCCMCCTCATCGAAACTCACCACCACCCCA 4000
35
                   NEHPGRIHLIETHUPH
               CA DECCAPTICOCCUTENCECAACTCACCACCTCCACCAACCCCACCTAC 4050
                 R L T N N T L H T P H L T P I T T
40
               ONCOR ORACOCOMORACOROCOCORROROSOCOROCOCORROCOCOR. 4150
                H H N T T T T P N T F F L N P N
               DOACGCCATCCTCATCACCGCCGCTCCGGCACCCTCGCCACCATCCTCG 4200
                  HAILITGGSGTLAGIL
              OCCGCOMCCTCAACCACCCCCACACCTACCTCCTCTCCCGCACACCACCA 4250
45
              A R H L N H P H T Y L L 3 R T P P
               PPTTPGTHIPCOLTDPT
               CCAAATCACUCAAGCCCTCACCCACATACCACAACCCCTCACCGGCATCT 4350
50
                  QITQALTHIPQPLTSI
              TOUNCACOGOGOCACCCTOGACGACGCCACCCTCACCAACCTCACCCCC 4400
              F H T A A T L D D A T L T B L T P
               CAACACCTCACCACCACCCTCCAACCCAAAGCCGACGCCTGGCACCT 4450
                 Q H L T T T L Q P K A D A A W H L
                CACCACCACACCACAAAACCAACCCCTCACCCACTTCGTCCTCTACTCCA 4500
55
                  H H H T Q N Q P L T H F V L Y S
                  S A A A T L G S P G Q A U Y A A A
              ARCSCOTTOOTOGROSSCOTOGOCACCOCACACCCAAGGRCAACC 4630
60
               NAFLDALATHRHTQGQP
```

PCT/US99/22886 WO 00/20601

DA COGA DAGOGA COGOGA POGOA TO DO COGOGOGO TILOTIG. 4740 it to both Both Both Both Both Both

The AvrII-Xhol hybrid FK-506 PKS module 8 containing the AT domain of module 12 of rapamycin is shown below.

```
10
    M F L Y E A A F F T G 3 P V T
        A A A L D D A P D V P L L A G L R
    GCGTA COMO COTOCOGO OTO COCOTOCOGO SA A O GCT CT CTOCOCO SA CO 150
      B T T V & B A A V R E R S D A D
    GCTGGCGGTGCTGCCGGMCGAGGAGGGGGGGGGACGCGTGCGTGTGTTTG 200
    R S P C C P T T S A P T P P S R S
    TOOTGGAACAGCACCGCCASCGTGCTCGGCCASCTGGGSCGCCGAAGADAT 250
     S W N S T A T V L G H L G A E D I
20
    COCCCCARCGACGACGATCAAGGAACTCGGCATCGACTCGCTCACCCCCA
     FATTFEELSIDSLTA
    TOCAGGTGGGGAAGGGGGTGACCACGGGGGAGGGGGTAGGGGTAAGGGG
     -Q L S N A L T T A T G V F L N A
    25
    TAVEDEPTPRALAARLG
    GGACGAGCTGGCCGGTACCCGCGCGCGCCCGTCGCGGCCCGGACCGCGCCA 450
     DELAGTRAPVAARTAA
    DOGGGGGGGGGAGGAAGGGGTGGGGATGGTGAGGCATGGCCTGCCGT 500
    T A A A H D E F L A I V G M A C R
30
    CTGCCGGGGGGGGTCGCGTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550
    L P G G V A S P Q E L W R L V A S
    CGSCACCGACGCATCACGGASTTCCCCGCGGGACCGCGGCTGGGACGTGG 600
      \texttt{G} \quad \texttt{T} \quad \texttt{D} \quad \texttt{A} \quad \texttt{I} \quad \texttt{T} \quad \texttt{E} \quad \texttt{F} \quad \texttt{P} \quad \texttt{A} \quad \texttt{D} \quad \texttt{R} \quad \texttt{G} \quad \texttt{W} \quad \texttt{D} \quad \texttt{V} 
    ACGCGCTCTACGACCCGGACCCCGACGCCATCGCCAAGACCTTCGTCCGG 650
35
    DALYDPDFDAIGKTFVR
    CACGGCGGCTTCCTCGACGGTGCGACCGGCTTCGACGCGGCGTTCTTCGG 700
    HGGELDGATSFDAAFFG
    GAT DAGCCCCCCCAAGGCCCTGECCATGGACCCGCAGCAACGGGTGCTCC 750
    I SPREALAM DPQQRV1
    TSGAGAGGTCCTGGGAGGCGTTCGAAAGCGCGGGCATCACCCCGGACGCG 800
40
    LETSWEAFESAGITPDA
    SCGCGGGGCAGCGACACCGGCGTGTTCATCGGGGGGTTCTCCTACGGGTA 850
    A R G S D T G V F I G A F S Y G Y
    CUGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTCGCAGACCA 900
45
    GTSADTNSFGATGSQT
    GCGTGC.CTCCGGCCGCCTCTCGTACTTCTACGUTCTGGAGGGCCCTTCG 950
    S V L S G R L S Y F Y G L E G P S
    STOACGGTCGACACCGCCTGSTCGTCGTCACTGGTCGCCCTGCACCAGGC 1900
    V T V D T A C S S S L V A L H Q A
5Û
    AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTCGGCGGGTG 1050
    G Q S L R S G E C S L A L V G G
    TCACGGTGATGGCGTCGCCGGGCGGGGGATTCGTCGAGTTCTCCCGGCAGCGC 1100
    V T V M A S P G G F V E F S R Q R
    55
    G L A P D G R A K A F G A G A D G
    TAGGAGGTTGGCGGAGGGGGGGGGGGTGGGTGGTGGTGGAGGGGGTGTGGG
     T S F A E G A G A L V V E R L S
    ACGCGGAGCGCCACGGCCACACCSTCCTCGCCCTCGTACGCGGCTCCGCG 1250
    D A E R H G R T V L A L V R G S A
```

	30TAA 0T 0 0 0A 0 3 0 0 3 0 3T 0 3W 23 3T 0T 0T 0T 0G 0 0 0 3A A 0 3 9 0 0 0 0 7	131
	00A 6 3AA 6 3 0AT 6AT 6AA 66A 6G9CCT 0G0GAA 7 30 3AAA 6T 6A 626CG	135
_		
5	10.88.091000.090390001.590908.0030808003008.0203001.0930580	140
	2001#03h33050h36536,30 <mark>0700000W00</mark> #h326h3h32h3033366h0	145
	10000000000000000000000000000000000000	150
1()		
	347273433,78723337,3873 77437	150
	ABBUAGIIKWWIAIPEG	
	SANOTACCARANANCTACHABAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	161
	E D P P T D H A D E F D P H T D W	
15	- BADBBTDBBTBCBTCBMBTTCBTGMBTDCDBBCBTBBCBTBBCBBAAAAA	165.
• •	C A G A V B L L T S A R F W C G	_0
	TOTAL TO A	170
		2 / 0
		175
20	AASGCCACCTCMTCCTGGAAAGCGCACCCCACTCAGCCTGCGGAAAA	175:
20	NARVILESAPPOOPPASN	200
	OGCGGTGATCGAGCGGCCACCGGAGTGGGTGCCGTTGGTGATTTCGGCCA	150
	A V I B B A P B W V P I V I S A	
	GGACCCACTCGGCTTTCACCGAGCACGACGACGCCGCTCCCCTCCCT	1851
A =	FTQSALTEHEGPLRAYL	
25	303306TCGCCGGGGGTGGATATGCGGGGCTGT3GCATCGACGCTGGGGAT	1991
	AASPGVOMRAVASTLAM	
	GACACGGTCGGTGTTCGAGCACCGTGCCGTGCTGCTGGGAGATGACACCG	1950
	TOROGGGACOGCTGTGTGTGACCCTCGGGGGGTGTTGGTGTTGCCGGGA	2000
30	V I G I A V S D P R A V F V E P G	
		2050
	CAGGGGTCGCABCGTGCTGGCATGGGTGAGGAACTGGCCGCCGCGCCTTCCC	2050
	CAGGGGTCGCAGCGTGCTGGCATGGGTGAGGAACTGGCCGCGGGGGTTCCC	2050
	CAGGGGTCGCAGCTGCTGGCATGGGTGAGGAACTGGCCGCCGCGTTCCC C G S Q R A G M G E E L A A A F P CGTCTTCGCGCGGATCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCG	
35	CAGGGGTCGCAGCGTGCTGGCATGGGTGAGGAACTGGCCGCGGCGCGTTCCC Q G S Q R A G M G E E L A A A F P CGTCTTCGCGCGGATCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCG V F A R I B Q Q V W D L L D V P	
35	CAGGGGTOGOAGOGTGCTGGCATGGGTGAGGAACTGGCCGCGCGCGTTCCC Q G S Q R A G M G E E L A A A F P OGTCTTCGGGGGATCCATCAGCAGGTGTGGGACCTGCTGGATGTGCCG V F A R I B Q Q V W D L L C V P ATCTGGAGGTGAACGAGACCGGTTACGCCCAGCCGGCCCTGTTUGCAATG	
35	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGGGGTTCCC Q G S Q R A G M G E E L A A F P CGTCTTCGGGGGATCCATCAGCAGGTGTGGGACCTGCTGGATGTGCCG V F A R I H Q Q V W D L L D V P ATCTGGAGGTGAACGAGACCGGTTACGCCCAGCCGGCCGTGTTUGCAATG E L E V N E T G Y A Q P A L F A M	2160
35	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGCGCG	
35	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGCGCG	2160
	CAGGGGTCGCASCGTGCTGGCATGGGTGAGGAACTGGCCGCGCGCGTGTTCCC Q G S Q R A G M G E E L A A A F P CGTCTTCGCGCGGGTCCATCAGCAGGTGTGGGACCTGCTGGATGTGCCCG V F A R I H Q Q V W D L L D V P ATCTGGAGGTGAACGAGACCGGTTACGCCCAGCCGGCCTGTTUGCAATG L L E V U E T G Y A Q P A L F A M CAGGTGGCTCTGTTCGGGCTGGAATCGTGGGGTGTACGACCGGACCC L V A L F G L L E S W G V R F D A GGTGATCGGCCATTCGGTGGGGTGACCTGCGGCTGCGTATGTGTCCGGGG	2160
35	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGCGCG	2160 2150 2200 2000
	CAGGGGTCGCASCGTGCTGGCATGGGTGAGGAACTGGCCGCGCGCGTGTTCCC C G S Q R A G M G E E L A A A F P CGTCTTCGCGCGGGATCCATCAGCAGGTGTGGGACCTGCTGCATGTGCCCG V F A R I B Q Q V W D L L D V P ATCTGGAGGTGAACGAGACCGGTTACGCCCAGCCGGCCCTGTTUGCAATG E L E V N E T G Y A Q P A L F A M CAGGTGGCTCTGTTCGGGCTGGAATCGTGGGGTGTACGACCGGACCC C V A L F G L L E S W G V R F D A GGTGATCGGCCATTCGGTGGGGTTGCGGCTGTATGTGTCCGGGG V I G H S V G E L A A A Y V S G TGTGGTCGTTGGGAGGATGCCTGCGGCTCGGGCTCGTCTCTCTC	2160 2150 2200 2000
	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGGGGTTTCCC C G S Q R A G M G E E L A A A F P CGTCTTCGGGGGTCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCG V F A R I B Q Q V W D L L D V P ATCTGGAGGTGAACGAGACGGTTACGCCCAGCCGGCCTGTTUGCAATG E L E V N E T G Y A Q P A L F A M CAGGTGGCTCTGTTCGGGCTGGAATCGTGGGGTTACGACCGGACGC L V A L F G L L E S W G V R F D A GGTGATCGGCCATTCGGTGGAGCTTGCGGCTGTATGTGTCCGGGG V I G H S V G E L A A A Y V S G TGTGGTCGTTGGAGGATGCCTGCACTTTGGTGTCGGGGCTCGTCTGTCGGGCTCGTCTGTCGGGGCTCGTCTGTCGGGGCTCGTCTGTCGGGCTCGTCTGTGTCGGGGCTCGTCTGTGTCGGGGCTCGTCTGTGTCGGGGCTCGTCTGTGTGTCGGGGCTCGTCTGTGTGTCGGGGCTCGTCTCTGTGTCGTGTCGGGGGCTCGTCTCTGTGTCGTGTCGGGGCTCGTCTCTGTGTCGTC	2160 2150 2200 2260 2300
	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGGGGTTTCCC C G S Q R A G M G E E L A A A F P CGTCTTCGGGGGATCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCG V F A R I B Q Q V W D L L D V P ATCTGGAGGTGAACGAGAGCTGTACGCCCAGCCGGCCCTGTTUGCAATG E L E V N E T G Y A Q P A L F A M CAGGTGGCTCTGTTCGGGCTGGAATCGTGGGGTGTACGACCGGACCC L V A L F G L L E S W G V R F D A GGTGATCGGCCATTCGGTGGGGTTGCGGCTGTATGTGTCCGGGG V I G H S V G E L A A A Y V S G TGTGGTCGTTGGGAGGCTGCACTTTGGTGTCGGGGGCTCGTCTGTCGGACCCGGACCCTGCGACCCTGGGACCCTGGGACCCTGGGACCCTGGGACCCTGGGGGGCTCGTCTGGGGCTCGTCTGGGGCTCGTCTGGGGCTCGTCTGGGGCTCGTCTGGGGCTCGTCG	2160 2150 2200 2260 2300
40	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGGGGTTCCCC C G S Q R A G M G E E L A A A F P CGTCTTCGGGGGATCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCG V F A R I B Q Q V W D L L D V P ATCTGGAGGTGAACGAGACCGGTTACGCCCAGCCGGCCCTGTTUGCAATG E L E V N E T G Y A Q P A L F A M CAGGTGGCTCTGTTCGGGCTGGAATCGTGGGGTGTACGACCGGACCC L V A L F G L L E S W G V R F D A GGTGATCGGCCATTCGGTGGGGTTGCGGCTGTATGTGTCCGGGG V I G H S V G E L A A A Y V S G TGTGGTCGTTGGGAGGTGCACTTTGGTGTCGGGGGCTCGTCTGTG V W S L E D A C T L V S A R A R L ATGCAGGCTCTGCCGGGGGTGGGATGGTCCCCGGTCTCGGA M Q A L P A G G V M V A V P V S E	2160 2150 2200 2050 2300 2350
	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGGGGTTCCCC C G S Q R A G M G E E L A A A F P CGTCTTCGGGGGATCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCG V F A R I B Q Q V W D L L D V P ATCTGGAGGTGAACGAGAGCGGGTTACGCCCAGCCGGCCG	2160 2150 2200 2050 2300 2350
40	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGGGGTTCCC C G S Q R A G M G E E L A A A F P CGTCTTCGGGGGTCATCAGCAGGTGTGGGACCTGCTCGATGTGCCG V F A R I B Q Q V W D L L D V P ATCTGGAGGTGAACGAGGTGTGAGCCCAGCCGGCCGGCTGTTUGCAATG L E V M E T G Y A Q P A L F A M CAGGTGGCTCTGTTCGGGCTGGAATCGTGGGGTGTACGACCGGACGC L V A L F G L L E S W G V R F D A GGTSATCGGCCATTCGGTGGAGCTTGCGGCTGTATGTGTCCGGGG V I G H S V G E L A A A Y V S G TGTGGTCGTTGGAGGATGCCTGCACTTTGGTGTCGGGGGGCTCGTCTG V W S L E D A C T L V S A R A R L ATGCAGGCTCTGCCGGGGGTGGGGTGGAGGTCGGGATCGGGA M Q A L P A G G V M V A V P W S E GGATGAGGCCCGGGCCGTGCTGGGGTGGAGGTCGCCGGGTCA C E A R A V L G E G V E I A A V	2160 2150 2200 2050 2300 2350 2400
40	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGGGGTTCCC G G S Q R A G M G E E L A A A F P OGTOTTCGGGGGATCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCG V F A R I B Q Q V W D L L D V P ATCTGGAGGTGAACGAGCTGTTAGGCCCAGCCGGCCTGTTUGCAATG L L E V M E T G Y A Q P A L F A M CAGGTGGCTGTTCGGGCTGGGAATCGTGGGGTGTACGGCCGGACCC L V A L F G L L E S W G V R F D A GGTGATCGGCAATCGGTGGGTTGCGGCTGGTATGTGCGGGG V I G H S V G E L A A A Y V S G TGTGGTCGTTGGAGGATGCCTGCACTTTGGTGTCGGGGGCTCGTCTG V W S L E D A C T L V S A R A R L ATGCAGGCTCTGCCCGGGTGGTGGTGGGATCGCCGGGTCTCGGA M Q A L P A G G V M V A V P V S E GGATGAGGCCCGGGCCGTGCTGGGTGGAGGTCGCCGGGTCAC C E A R A V L G E G V E I A A V ACGGCCCGTCGTCGGGTGGTTCTCCCGGTGATGAGGCCLLIGGTGCAGG	2160 2150 2200 2050 2300 2350 2400
40	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGGGGTTCCC G G S Q R A G M G E E L A A A F P OGTOTTCGGGGGATCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCG V F A R I B Q Q V W D L L D V P ATCTGGAGGTGAACGAGGTGTACGCCCAGCCGGCCGGTGTUGCAATG E L E V M E T G Y A Q P A L F A M CAGGTGGCTCTGTTCGGGCTGGAATCGTGGGGTGTACGACCGGACCG C V A L F G L L E S W G V R F D A GGTSATCGGCCATCCGGTGGAGCTTGCGGCTGTATGTGTCCGGGG V I G H S V G E L A A A Y V S G TGTGGTCGTTGGAGGATCCCTGCACTTTGGTGTCGGGGGGCTCGTCTG V W S L E D A C T L V S A R A R L ATCCAGGCTCTGCCGGGGGGGGGTGGTGGGATCGCCGGTCTCGGA M Q A L P A G G V M V A V P V S E GGATGAGGCCCGGGCCGTGCTGGGTGGGGGGGCTCGGCCA C E A R A V L G E G V E I A A V ACGGCCCGTCGTGGGTGGTTGCTCCCGGTGAGGGCLLIGGTGCAGG M G P S S V V L S G D E A A V L Q	2150 2150 2200 2250 2350 2350 2400 2450
40	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGGGGTTCCC G G S Q R A G M G E E L A A A F P OGTOTTCGGGGGATCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCG V F A R I B Q Q V W D L L D V P ATCTGGAGGTGAACGAGGTGTAGGCCCAGCCGGCCTGTTUGCAATG L L E V M E T G Y A Q P A L F A M CAGGTGGCTGTTCGGGCTGGAATCGTGGGGTGTACGACCGGACCG L V A L F G L L E S W G V R F D A GGTSATCGGCCATCGGTGGACTTGCGGCTGTATGTGTCCGGGG V I G H S V G E L A A A Y V S G TGTGGTCGTTGGAGGATCCTTGGTGTCGGCGGGGCTCGTCTG V W S L E D A C T L V S A R A R L ATCCAGGCTCTGCCGGGGTGGTGGGTGGGATCGCCGGGTCTCGGA M Q A L P A G G V M V A V P V S E GGATGAGGCCCGGGCCGGTGGTGGGTGGAGGATCGCCGGGTCA C E A R A V L G E G V E I A A V ACGGCCCGTCGTGGGTGGTGGGTGATGAGGCCLIGGTGCAGG M G P S S V V L S G D E A A V L Q GCCGCGGAGGGGCTGGGGGAAGGGGGTGGGGACCAGCCAG	2150 2150 2200 2250 2350 2350 2400 2450
40	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGGGGTTCCC G G S Q R A G M G E E L A A A F P OGTOTTCGGGGGATCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCG V F A R I B Q Q V W D L L D V P ATCTGGAGGTGAACGAGGTGTAGGCCCAGCCGGCCTGTTUGCAATG L L E V M E T G Y A Q P A L F A M CAGGTGGCTGTTCGGGGTGGGAATCGTGGGGTGTACGGCCGGGCCGGCC	2150 2150 2200 2250 2300 2350 2400 2450
40	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGGGGTTCCC G G S Q R A G M G E E L A A A F P OGTOTTCGGGGGATCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCG V F A R I B Q Q V W D L L D V P ATCTGGAGGTGAACGAGGTGTAGGCCCAGCCGGCCTGTTUGCAATG L L E V M E T G Y A Q P A L F A M CAGGTGGCTGTTCGGGCTGGAATCGTGGGGTGTACGACCGGACCG L V A L F G L L E S W G V R F D A GGTSATCGGCCATCGGTGGACTTGCGGCTGTATGTGTCCGGGG V I G H S V G E L A A A Y V S G TGTGGTCGTTGGAGGATCCTTGGTGTCGGCGGGGCTCGTCTG V W S L E D A C T L V S A R A R L ATCCAGGCTCTGCCGGGGTGGTGGGTGGGATCGCCGGGTCTCGGA M Q A L P A G G V M V A V P V S E GGATGAGGCCCGGGCCGGTGGTGGGTGGAGGATCGCCGGGTCA C E A R A V L G E G V E I A A V ACGGCCCGTCGTGGGTGGTGGGTGATGAGGCCLIGGTGCAGG M G P S S V V L S G D E A A V L Q GCCGCGGAGGGGCTGGGGGAAGGGGGTGGGGACCAGCCAG	2150 2150 2200 2250 2300 2350 2400 2450
40	CAGGGGTCGCASCGTGCTGGCATGGGTGAGGAACTGGCCGCGGGGTTCCC G G S Q R A G M G E E L A A A F P OGTOTTCGCGGGGATCCATCAGCAGGTGTGGGACCTGCTGGATGTGCCCG V F A R I B Q Q V W D L L C V P ATCTGGAGGTGAACGAGGTGTAGGCCCAGCCGGCCTGCTUGCAATG E L E V N E T G Y A Q P A L F A M CAGGTGGCTCTGTTCGGGCTGGAATCGTGGGGTTACGCCGGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGCCGACCGGACCGCCCGCCCTGTATGTGTCCGGGGCT C V A L F G L L E S W G V R F D A G GGTGATCGGCGATTCGGGGTTGGGGGCTGGTATGTGTCCGGGGCTCGTCT V W S L E D A C T L V S A R A B L ATGCAGGCCTGGCGGGCTGGGGTGATGGTCGCGGTCTCGGA M Q A L P A G G V M V A V P V S E GGATGAGGCCCGGGGCTGGGTGAGGGTGGGGGCTCGCGGTCA C E A R A V L G E G V E I A A V A CGGCCCGGGGCTGGGGGGGGGGGGGGGGGGGGGGGGG	2150 2150 2200 2250 2300 2350 2400 2450
40	CAGGGGTCGCASCGTGGCATGGCTGAGGAACTGGCCGCGCGCGTTTCCC G G S Q R A G M G E E L A A A F P OGTOTTCGCGGGGATCCATCAGCAGGTGTGGGACCTGCTGGATGTGCCCG V F A R I B Q Q V W D L L C V P ATCTGGAGGTGAACGAGGTGTAGGCCCAGCCGGCCTGCTUGCAATG E L E V N E T G Y A Q P A L F A M CAGGTGGCTCTGTTCGGGCTGGAATCGTGGGGTTAACGCCGGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGCCCGACCGGACCGGCCGGACCGCCCGGACCGCCCGATTCGGGGTTATGTGTCCCGGGTCATTGTGTGTCGGGGGCTCGTCTCGGA W S L E D A C T L V S A R A B L ATGCAGGCTCTGCCGGGGTGGAGGGTGGGGGACCAGCCGGGTCA N Q A L P A G G V M V A V P V S E GGATGAGGCCCGGGCGGTGGGTGAGGGTTGGAGATCGCCGGGTCA C E A R A V L G E G V E I A A V ACGGCCCGGGCCGGTGGTGGGGGGGTGGGGGACCAGCCACCGTT A A E B L G K W T R L A T S H A F CCATTCCGCCCGTTATGGAACCCATGCTGGAGGACCAGCCGCGCGTCGCCG H S A P M E F M L E E F P A V A	2150 2150 2200 2250 2300 2350 2400 2500 2500
40	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGGGGTTCCC G	2150 2150 2200 2250 2300 2350 2400 2500 2500
40	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGGGGTTCCC G	2150 2150 2200 2250 2300 2350 2400 2500 2500
40 45 50	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGCGTTCCCCCCCC	2150 2150 2200 2250 2300 2350 2400 2500 2500
40 45 50	CAGGGGTCGCAGGGTGGCATGGGTGAGGAACTGGCCGCGCGTTGCCCCCCCC	2150 2150 2200 2250 2300 2350 2400 2500 2500 2600
40 45 50	CAGGGGTCGCASCGTGCTGGCATGGGTGAGGAACTGSCCGCCSCCTTCCC C G S Q R A G M G E E L A A A F P CGTCTTCGCSCGGATCCATCAGCAGGTGTGGGACCTGCTGGATGTSCCCG V F A R I B Q Q V W D L L D V P ATCTGGAGGTGAACGAGAGCGGGTGAGCGGGGCCGTGTTUGCAATG E L E V M E T G Y A Q P A L F A M CAGGTGGCTGTTCGGGGTGGGAATCGTGGGGTGTACGGCGGGACGC L V A L F G L L E S W G V R F D A GGTSATCGGCCATTCGGTGGGTGAGCTTGCGGCTGGTATGTGTCCGGGG V I G H S V G E L A A A Y V S G TGTGGTCGTTGGAGGGTGACTTTGGTGTGGGGGGTCTCTG V W S L E D A C T L V S A R A R L ATGCAGGCTCTGCCGGGGTGGGTGAGGTGGGGAGGCCGGGTCTCGGA M Q A L P A G G V M V A V P V S E GGATGAGGCCGGGGTGGTGGTGGGGTGGGGAGAGCCGGGTCA C E A R A V L G E G V E I A A V ACGGCCCGTGGTGGTTGTCTCCCGGTGATGGGGGACAGGGCCGGGTTA M G P S S V Y L S G D E A A V L Q GCCGCGGGGGGGTGGGGGGGGGGGGGGGGGGGGGGGTTA A E G L G K W T R L A T S H A F CCATTCCGCCCGTAGGGGAGGGGGGGGGGGGGGGGGGG	2150 2150 2200 2250 2300 2350 2400 2500 2500 2600
40 45 50	CAGGGGTCGCASCGTGCTGGCATGGGTGAGGAACTGSCCGCCSCCTTCCC C G S Q R A G M G E E L A A A F P CGTCTTCGCSCGGATCCATCAGCAGGTGTGGGACCTGCTGGATGTSCCCG V F A R I B Q Q V W D L L D V P ATCTGGAGGTGAACGAGACCGGTTACGCCCAGCCGGCCCTGTTUBCAATG E L E V M E T G Y A Q P A L F A M CAGGTGGCTCTGTTCGGGCTGGGAATCGTGGGGTGTACGCCCGGACGC L V A L F G L L E S W S V R P D A GGTSATCGGCCATTCGGTGGATTGGGGCTGGTATGTGTCCGGGG V I G H S V G E L A A A Y V S S TGTGGTCGTTGGAGGGTGCACTTTGGTGTGGGGGGTCGTCTGT V W S L E D A C T L V S A R A R L ATGCAGGCTCTGCCGGGGGTGGGTGATGGTGGGGGGTCGGGACA M Q A L P A G G V M V A V P V S E GGATGAGGCCCGGGGTGGTGTGGGGTGAGGTTGGGAGATCGCCGGGTCA C E A R A V L G E G V E I A A V ACGGCCCGGGGTGGTGTTCTCTCCGGGTGATGAGGTCGCGGGTCA C E A R A V L G E G V E I A A V ACGGCCCGGGGTGGTGTTCTCTCCGGGTGATGAGGCLLLGGTGCGAG M G P S S V V L S G D E A A V L Q GCCGGGGAGGGGGTGGGGGAGAGCGGGTGGGGGACCAGCAGCAGCAGC H S A P M E F M L E E F P A V A AAGGCCTGACCGGTGGTGGTGGCGGGGGGGGGGGGGG	2100 2150 2200 2250 2300 2350 2400 2500 2500 2600 2700
40 45 50	CAGGGGTCGCASCGTGCTGGCATGGGTGAGGAACTGSCCGCCSCCTTCCC C G S Q R A G M G E E L A A A F P CGTCTTCGCSCGGATCCATCAGCAGGTGTGGGACCTGCTGGATGTSCCCG V F A R I B Q Q V W D L L D V P ATCTGGAGGTGAACGAGAGCGGGTGAGCGGGGCCGTGTTUGCAATG E L E V M E T G Y A Q P A L F A M CAGGTGGCTGTTCGGGGTGGGAATCGTGGGGTGTACGGCGGGACGC L V A L F G L L E S W G V R F D A GGTSATCGGCCATTCGGTGGGTGAGCTTGCGGCTGGTATGTGTCCGGGG V I G H S V G E L A A A Y V S G TGTGGTCGTTGGAGGGTGACTTTGGTGTGGGGGGTCTCTG V W S L E D A C T L V S A R A R L ATGCAGGCTCTGCCGGGGTGGGTGAGGTGGGGAGGCCGGGTCTCGGA M Q A L P A G G V M V A V P V S E GGATGAGGCCGGGGTGGTGGTGGGGTGGGGAGAGCCGGGTCA C E A R A V L G E G V E I A A V ACGGCCCGTGGTGGTTGTCTCCCGGTGATGGGGGACAGGGCCGGGTTA M G P S S V Y L S G D E A A V L Q GCCGCGGGGGGGTGGGGGGGGGGGGGGGGGGGGGGGTTA A E G L G K W T R L A T S H A F CCATTCCGCCCGTAGGGGAGGGGGGGGGGGGGGGGGGG	2100 2150 2200 2250 2300 2350 2400 2500 2500 2600 2700

WO 00/20601 PCT/US99/22886

```
Chere i la a i gaba bah bay wa
                        TAGGGTGGAGTGGGGGGGGGATGGTGGGGGAACAG 2850
                          or the standard and a few and
          . 3080 Jan 38. john jakokan Adago prodesika da sakan da sakan sakan sakan da sakan da sakan da sakan da sakan d
          R V L L L F T Y A F Q H D R Y W L SASTOSSOTOSSOAC 2950
            E S A B B A T A B S G H B V L
           S V A V A S S F S R V F T S F
10
          FAGATEAVETAELALAF
           gadamagaan oo sharaagaah <mark>haadaraannan saraan</mark> aarah arraasi 3116
          15
            . For the contraction of the co
          ANISOTA DO BODEA DE BASE DO BODO CONTRA COSTO DA CARCO COSTO 330 - 3200
          SACGOODOGTGGACGCTGCACGCCGACGGGGTTCTCCGGCCCCGGCCGCGCGT 3250
           DAPWTLHAEGVLRP3RV
20
              PADGLPGAWRRADQUEI
           SAAGOSSAAGTOSAGAGOOTTGACGGOTTCGTGGCACACCCCGACCTSCT 3400
25
           E A E V S S P D G F V A H P C L
          DGADGDGGTCTTCTDCGCGGTCGGGGACGGGAGCCGCCAGCCGACCGGAT 3450
             DAVESAVGDGSRQPT3
          DECENERACIONES CENTRO DE CONTRO DE C
          W P I L A V H A S D A T V L R A C
30
          STORCOGGGGGGACKGTGGTGTGGTGGAGCTCGCCGTTCGACGGTGC 3550
           L T R R D S G V V E L A A F D G A
          CGGAATGCCGGTGCTCACCGCGGGAGTCGGTGACGCTGGGGGAGGTCGCGT 3600
            G M F V L T A E S V T L G E V A
          DGGCAGGCGATTCGGACGACGACGGTCTGCTTCGGCTTGAGTGGTTG 3650
35
          S A S S S E S D G L L R L E W .
          00FE ATTREEPARPERPARPARE ADDOOR THE RESERVE ATTREEPARPER OF TO DOOR TO DOOR
           E M A E A H M D G A C E L P E G M
          CACCCTCATCACCGCCACACACCCCGACGACCCCGACGACCCCACCAACC 3750
            40
          CCCACACACACACCCACACCCACACACACACACACGCGTCCTCACC 3800
          PHHTPTRTHTQTTRVLT
          GCCCTCCAACACCACCTCATCACCACCAACCACCACCTCATCGTCCACAC 3850
           ALQBELITTNETLIVET
          CACCACCGACCCCCAGGCGCCGCCGTCACCGGCCTCACCCGCACCGCACCGCAC 3900
45
            T T D P P G A A V T G L T R T A
         AAAAGGAACAGGGGGATCCACCTCATCGAAAGGCACCACGGGGAC 3950
          O N E H F G R I H L I E T H H F H
         ACCCCACTCCCCCTCACCCAACCCACCCTCCACCAACCCCACCTACG 4000
50
           T F L F L T Q L T T L H Q F H L P
         ACCACACACACACCACAACACCCCCAACACCCCCACCCCTCAACCCCAAC 4100
          H H H T T T T P N T F P L M P H
          CACGCCATCCTCATCACCGGCGGGCTCCGGCACCCTCGCCGGCATCCTCGC 4150
55
           HAILITGGSGTLAGILA
          COSCOMOCT CARCOACCCCACACCTACCTCCTCTCCCGCACACCACCAC
             R R L M R P R T Y L L S R T F P
          COCCCACACACCCGGCACCCCACATCCCCTGCGACCTCACCGACCCCACC 4250
          PPT PSTHIPCDLT DPT
60
```

WO 00/20601 PCT/US99/22886

The AvrII-AhoI hybrid FK-506 PKS module 8 containing the AT domain of module 13 of rapamycin is shown below.

```
30AT83330T8TX8GAGGGGGAGCAGGGGAACGGGAACT33CTGGTGGTG
     BELYEAARRTGSPVVV
   25
   A A A L D D A P D V P L L R G L R
   GOGTACGMOCOTTCGGGGGTGCCGCCGTCCGGGGAACGCTCTCTCGCCGACC 150
    A T T V R R A A V R E R S L A D
   RSPCCPTTSAPTPPSRS
   TOSTGGAACAGCACOGCCACOGTGCTCGGCCACCTGGGCGCCGAAGACAT 250
    SWNSTATVLGHLGAED
   CCCGGCTACGACGACGTTCAAGGAACTCGGCATCGACTCGCTCACCGUGG 300
    FATTTFRELGIDSLTA
   TOURGOTGOOGHADGOGOTGROCROGGGGADOGGGGTRAGGGTCARCGCC 380
35
    QLASALTTATGVRLNA
   A0A3033T0TT03A3TTT033A03C3C3C3C3C3C3C3C3C3C3C3A3A3TC3C 400
    TAVES PTERALAARUS
   CARCAMATTABOOGGTACOCACACACACACACACACACACACACACA
    DELAGTRAPVAARTAA
40
   COGCOGCOCCACGACGAACCCCTGGCGATCGTGGGCATGGCCTGCCGT 500
   TAAAHDEPLAIVGMACR
   CTGCCGGGGGGGTCGCGTCGCCACAGJAGCTGTGGCGTCTCGTCGCGTC 550
    L P G G V A S P Q E L W R L V A S
   CGGCACCGACGCATCACGGAGTTCCCCGCGGACCGGCGGCTGGGACGTGG 600
45
    G T D A I T E F P A D R G W D V
   ACGCCCTACGACCCGGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650
   DALYDPDPDAIGKTFVR
   CACGSCGGCTTCCTCGACGGTGCGACGGGCTTCGACGCGGGTTCTTCGG 700
   H G G E L D G A T G F D A A F F G
50
   GATCAGCCGGGGGGGGCGTGGCGATGGAGCGAGCAACGGGTGCTCC 750
    I SPREALAM DPQQRVL
   TGGAGACGTCCTGGGAGGCGTTCGAAAAGCGCGGGCATCACCCCCGGACGCG 800
   leteweares a gireba
   SOSCESSCASSCACACOGOCCTSTTCATCGOCCGCTTCCCCTACGGGTA 850
55
   AFGSSTEVELSAFSYGY
   G T G A D T N G F G A T G S Q T
   SCGTGCTCTCCSGCCGCCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950
   SVISGRISYFYGLEGPS
```

```
AGRICA GTOCCTOCCCTOGGGCAMTGCTCGCTCGCTCGCTGGCGCGTG (113)
                                       \texttt{Intitude}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppose}(\mathsf{Suppos
                     G L A P D G R A K A F G A G A C
                  10
                     TOFAEGAGALUVEA:
                 ACBOSASOSCACESCOACACCETCOTOSCCCTACECESCUTUCOCE 1250
                  GCTMICTCCGMCGGCGGGTGGAAAGGGTGTGTGGGGGGGGGAAAGGGGCGCTGT 1300
                   A N S C S A S N G C S A F N S F
 15
                  CORGURACOURTOR CORGODOTO CORDO PROTERO TOR CONTROLO DE                     TO THE REPORT OF A CONTROL OF THE PROPERTY OF A CONTROL OF THE PROPERTY OF A CONTROL OF THE PROPERTY OF THE PR
                 A D V D A V E A H G T G T R L G
                 200ATCGNGBBGCRGSCCCTGCTCACCGACGRAGGACAGGACCCCCCCCCA
 20
                 fillogsbkswighapa
                   A S S V A S I I K M V Q A I R B S
25
                 SAACTSCCGCGACACTGCACGCGGACGAGCCGTCGCCGCACGTCGACTC 1800
                  E L P P T L H A D E P S P H T E W
                GACGGCCGGTCGAGCTCCTGACGTCGGCCCGGCCGTGGCCGGGGA 1650
                     TAGAVELLTSARPWF3
                 COUGTOGCCCTAGGCGGGCGGGCGTGTCGTCCTTCGGAGTCAGCGGCACC
30
                TGRPRRAGVSSFGVSG
                AACGCCCACGTCATCCTGGAGAGCGCACCCCCCGCTCAGCCCGCGGGGAAGGA 1750
                 NAHVILESAPPAQPAEE
                SGCGCAGCCTGTTGAGACGCCGGTGGTGGCCTCGGATGTGCTGCCCCTGG 1800
                     AQPVETPVVASDVLF
35
                TOATATOGOCCAAGACCCAGCCCCCCTGACCGAACACCGAAGACCGGGCTG 1850
               VISAKTOPALTEREPE
                                   FAYLAASPGACIAAVAS
                SACCOMBODGOMOROSOTOGOTOTTGARCACOCOCCOTROTCOTTC 1983
40
                   I LAVTRS V PEHRAVI
                GAGATGACACOGTCACOGGCACCGGGTGACCGACCCAGGATCGTGTTT 2000
                3 D D T V T G T A V T D P R I V F
                STOTTTOOCGGGCAGGGGTGGCAGTGGCTGGGGATGGGCAGTGCADTGDG 2050
                  V F P G Q G W Q W L G M G S A L R
45
                CGATTCGTCGGTGGTGTTCGCCGAGCGGATGGCCGAGTGTGCGGCGGCGCGT 2100
                   ESSVVEAERMAECAAA
               TGCGCGAGTTCGTGGACTGGGATCTGTTCACGGTTCTGGATGATCCGGCG 2150
                  . REFVDW D L F T V L D D P A
                STOSTGONOCGGGTTGATGTGGTCCAGCCCGCTTCCTGGGCGATGATGGT 2200
50
                  T T D R T D T V Q P A S W A M M T
                   S L A A V W Q A A G V B P D A 1
                TOBBOCATTOSCAGGGTGAGATOSCOGGAGCTTOTGTGGCGGGTGCGGTG 2306
                   GHSQGEIAAACVAGAV
55
                TOROTROGOGRITGOGGGGGGGATOGTGROCTTGDGGRGOCAGGGGRATOGG 2380
                  SLRDAARIVTLESQAIA
                3 0 1 A G R G A M A S t A L P A
               AGGATGTCGAGCTGGTCGACGGGGCCTGGATCGCCGCCACAACGGGCCC 2450
60
                1 D V E L V D G A W I A A H N G P
```

	PORT TO A CORDINATE BEST SOLD CARROUND AND CONTROL CONTROL CONTRACTOR CONTRACTOR CONTRACTOR CONTRACTOR CONTROL	250
5	TO THE TO A TO A TO BOTH BOOK BOTH AND THE TO THE TO THE TOTAL AND THE T	1600
	A 7 B T B H V B L F R C E L L C C C ANTAGRACAGACAGACAGACAGACAGACAGACAGACAGACAGA	2680
10	STANDSPORTED TO GREAT OF A CONTROL OF A CONT	2700
	Aditisogtigaadoggtoggtttogacocoggoostoagoaagtoagtoaggoo	2750
, ~	CASGGCGACACCSTGTTCGTCGAGGTCAGCGCCAGCCGGTGTTGTTGCA	
15	Garden with the American Cacacamero Accacamental and Cacacamental and Caca	2850
	GOGAUGGGAGGGGAGAGGGGCCTATGTCCACGGG S D A T S M L T A L A Q A Y V R G	2900
20	STOA DOSTOGA OTEGOOGOCATOCTOGGCACCADERCACCOGGGTACT V T V 5 W P A 1 L 3 T T T T R V L GGA DOTTOGGA COTA COCOTTOCAA CACCAGGGGTACT GGCTT GGAGT CG	2950
	SAMPO TEURNO TROQUOTTEONNUNCONSURATARE RAPERSAMENTA DE LE RECENTA DE RECENTA	2053
25		3100
	A V A G S P G R V F T G P V F A G	3150
	A D R A V F I A E L A L A A A D CCACCGACTGCCCCACGGTCGACGTCGACGTCACGTCGCCCGGC	3200
30	2011 66366670006110063311006441101104	3250
		3300
35	CGCCGACGGGGGGCGCCTTCACCGTCCACACCCGCGTCGGGGGACGCCCAAAAAAAA	3300 3350
35	CGCCGACGGGGGGCGCCTCACCGTCCACACCCGCGTCGGGGGACGCCCAAAAAAAA	
35	CGCCGACGGGCGCGCTTCACCGTCCACACCCGCGTCGGCGACGCCCCACACCCGCGTCGGCGACGCCCCACACCCGCGCGCG	3350
	CGCCGACGGGCGCGCTTCACCGTCCACACCCGCGTCGGCGACGCCC A D G R R R F T V H T R V G D A CGTGGACGCTGCACGCGGAGGGGGTTCTCGGCCCGGGCGGCGGCGCCCAG P W T L H A E G V L R P G R V P Q CCGGAAGCGGTGACACGGCGGCGGCGGGGGGGGGGGGG	3350 3400 3450 3500
40	CGCCGACGGGCGCCGCTTCACCGTCCACACCCGCGTCGGCGACGCCCAAAAAAAA	3350 3400 3450 3500 3550
40	CGCCGACGGGCGCCGCTTCACCGTCCACACCCGCGTCGGCGACGCCC A D G R R R F T V H T R V G D A CGTGGACGCTGCACGCGGAGGGGGTTCTCCGCCCGGCCGG	3350 3400 3450 3500 3550 3600
40	CGCCGACGGGCGCCGCTTCACCGTCCACACCCGCGTCGGCGACGCCCAAAAAAAA	3350 3400 3450 3500 3550 3600 3650
40	CGCCGACGGGCGCCGCTTCACCGTCCACACCCGCGTCGGCGACGCCC A D G R R R F T V H T R V G D A CGTGGACGCTGCACGCGGAGGGGGTTCTCCGCCCGGGCGGG	3350 3400 3450 3500 3550 3650 3700
40	CGCCGACGGGCGCCGCTTCACCGTCCACACCCGCGTCGGCGACGCCC A D G R R R F T V H T R V G D A CGTGGACGCTGCACGCGGGGGGTTCTCCGGCCCGGCC	3350 3400 3450 3500 3550 3650 3700 3750
40 45 50	CGCCGACGGGCGCCGCTTCACCGTCCACACCCGCGTCGGCGACGCCCAAAAAAAA	3350 3400 3450 3500 3550 3650 3700 3750
40	CGCCGACGGGCGCCGCTTCACCGTCCACACCCGCGTCGGCGACGCCCAAAAAAAA	3350 3400 3450 3500 3550 3650 3700 3750 3800
40 45 50	CGCCGACGGGCGCCGCTTCACCGTCCACACCCGCGTCGGCGACGCCCAAAAAAAA	3350 3400 3450 3500 3550 3600 3750 3750 3800 3850 3900

```
GA COCTOCARGOS CORTOTOA COESCOTOA COCTOA COGOA CARAA CAA 4001
| Color Postoa Aoa to to to to Color Fotoa (Color E
                               TO DESCRIPTION OF A STREET OF 
10
                            NUCTOA NON E COU TO GOTOCO GOA CACOTO GOLGO CATO CUTO GUE E COCACIA (A 2.5 C
                       omoning and open and omneemed for coccan and an open coccan a
                     OCCARGOCOTOROCOACATACOACAAGOCOTOROGGGCATOTTCCACAGO 4400
                 T Q A L T H I P Q P L T G I F H T
                 TO A A A DOCO DO A OTO A A DO A DO A DO BOA DO A DO BOA DA DO BOBO DA DA DO BOBO DA DA DO BOA DO BOBO DA DO BO
                  AATLEDATLINLPIE
20
                 CACCACCACCCTCCAACCCMMGCCGACGCCTGGCACCTCCAACCACC
                       TTUDERADAAXELEE
                 нтоморгонгистазаа
                GOCACCTCGGCAGCCCGGCCAAGCCAACTWCGCGGCCGCCAACGCCTT 4600
25
                  ATLSSPGQAUTAAAUAE
                 COTOGREGOCOTOGOCACCCACOGOCACACOGARGA DAR COCESCACOA 4600
                      LDALATHRHTQGQPAT
                 CCATOGOOTGGGGGATGTSSCACACCACCACACTCACCAGGGGAACTC 4700
30
                 TIAWGMWHTTTTLTSQL
                 ACCGACAGCGACCGCGACCGCATCCGCCGCGGCGGCTTCCTGCCGATCTC 4750
                  TOSORDRIRRGGFLPIS
                 GGACGACGAGGGCATGC
                        DDEGM
35
```

The *NheI-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module 12 of rapamycin is shown below.

```
GOATGOGGCTGTACGAGGOGGCACGGCGCACGGAAGTCCCGTGGTGGTG 50
     M R L Y E A A R R T G S P V V V
   40
   A A A L D D A P D V P L L R G L R
   RTTVRRAAVRERSLAS
   GOTOGOOGTGCTGCCGACGACGACGACGCGCCGACGCCTCCCTTCGC200
   R S P C C P T T S A P T P P S P S
   TOOTGGAACAGCACCGCCACCGTGCTCGGCCACCTG3GCGCGGAAGACAT 250
    SWNSTATVLGHUGAEDI
   CCCGGCGACGACGACGTTCAAGGAACTCGGCATCGACTCGCTCACCGCGG 300
    PATTFKELSIDS LTA
   TOCASCTGCGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGCC 350
50
   V Q L R N A L T T A T G V R L N A
   ACAGOGGTOTTOGACTTTCCGACGCCGCGCGCGCTCSCCGCGAGACTCGG 400
   TAVFDFPTPRALAARLG
   CGACGAGCTGGCCGGTACCCGCGCGCCCGTCGCGGCCCGGACCGCGGCCA 450
    DELAGTRAPVAARTAA
55
   nancaecogogoacghagahgaeragacakrogragaaAFgaaaTgaaaT
   T A A A H D E F L A I V 3 M A C R
   CTGCCGGGGGGGGTGGGGTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550
    L P G G V A S P Q E L W B L V A S
```

	UPBRACCEACGCATCACGGAGTCCCCCCGGGGGGGGGGCTGGGACGTGG	000
	- ACCCCTOTACCA COCCCA COCCCA COCCATOR CAACA COTTOCT COCC - ACCCCTOTACCA COCCCATOR CAACA COCCATOR CAACA CAACA COCCATOR CAACA CAACA CAACA CAACA CAACA CAACA CAACA CAACA C	9 650
5	- 1A 0000 a componed A plant à de A pode ont 195A 1900 a componed () - Hill 2 - Pill 1 - Dill 3 - A - Till 0 - Fill 1 - A - A - Fill Fill 3	
	- GATOAGCOCAGOGAGACCOCAGACCAGACCAGCAACCAGCAACCAGCACCACACACCAC	750
10	TGGAGACGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	* 100
	GPGCGGGCACACCGGCGTGTTCATCGGGGTTCTCCTACGGGTA A R G S C T G V F 1 3 A F S Y S Y	
	- DBSCACSSCTOCSSATACCAR CGSCTTCSSCCSCBACAGGGCCCAAAAAAAAAAAAAAAAAAAAAAA	. 7UL
15	GTGTGGTGTGGGGGGGGGGGTGTGGTAGTTGTAGGGGGGG	950
	STCACGGTCGACACCGCCTGCTCGTCGTCACTGGTCGCCCTGCACCAGGC	1000
20	AGGGCAGTCCCTGCGCTCGGGCGAATGCTCSCTCSCCCTGSTCGSCCSTS G Q S L R S G E C S L A L V G S	1050
-0	G Q S L R S G E C S L A L V G S TOACGGTGATGGTGTGGGGGGGGGGGTTGGTGGGGGGGGGG	1100
	GECAGOCOCORRO DE A KARE O A CARE DE A CORENTA DE CORENT	1150
25	TACGASCTTCOCCGAGGGCGCCGGTGCCCTGGTGGTGGAGGGGCTCTCCG	1200
	ACGCGGAGCGCCACACCACCACCGCCCCCCCCCCCCCC	
30	GOTAACTCOGACGGCGCGTCGAACGGTCTGTCGGCGCAACGGCCCCTC A N S D G A S N G L S A F N G P S	1300
	CCAGGAACGCGTCATCCACCAGGCCTTCGCGAACGCGGAAACTCACCCCG Q E R V I H Q A L A N A K L T F CCGATGTCGACGCGGTCGAGGCGCACGGCACCGGCTCGGCGAC	1400
	A D V D A V E A H G T 3 T R L G D	1100
35	CCCATCGAGGCGCAGGCGTGCTCGCGACGTACGGACAGGACCGGGCGAC P I E A Q A L L A T Y G Q D R A T	1450
÷	GCCCCTGCTGGGCTCGCTGAAGTCGAACATCGGGCACGCCCAGGCCG	
40	CSTCAGGGTCGCCGGGATCATCARGATGGTGCAGGCCATCCGGCAGGGGA	1880
	SAACTGCCGCCGACACTGCACGCGGACGAGCCGTCGCCGCACGTCGACTG	.000
	GACGGCCSTGCCGTCGACCTCGTGACCTGGCCGGCCGTGGCCGGGGA T A G A V E L L T S A R P W P G	1000
45	CCGGTCGCCGCGCGCGCGCTGCCGTCTCGTCGTTCGGCGTGAGCGGCACG	
	AACGCCCACATCATCCTTGAGGCAGGACCGGTCAAAACGGGACCGGTCGA $\mathbb N$ A $\mathbb N$ I I L $\mathbb E$ A S P V K T S P V $\mathbb E$	1750
50	AGATEAGPVEVSPVEA	1800
	GACOGOTOCOCGOGGGGGGGGGGGGGGGGGGGGGGGGGGG	
55	CTOGTGTOGGOGGTTOCCOGGAGGCACTCGACGAGCAGATCGGGCGCCTLVSARRSPEALDEQISEQISGCGTGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	
	R A Y L D T G P G V D R A A V A AGACACTGCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	
	2 T L A R R D H F T H R A V L L G GACACCGTCATCGGCGCCCCCCCGGGGACCA3GCCGAACTCGTCTT	
60	DTV;GAPPADÇADELVF	

	COTOTACTOSSETCASSSCAPCCAGOATOCOGATOSSCARGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	
5	- 20 90 00 23 10 00 02 3 10 10 3 20 39A	
	LOUNCE TO BOUND BOTH BOTH AND PORT AND PORT AND PORT AND PORT OF THE PROPERTY	1150
10	TACGACCGGACCGCCCCCCCCCCCCCCCCCCCCCCCCCC	2300
	TATSTSTOCCSSSTETSSTCSTTSSAGGATSSCTSCACTTGGTSTCSSC TATSTSTCCSSSSSTETSSTCSTTSSAGGATSSCTSCACTTGGTSTCSSC TOUR SOLO TOUR SOLO BOOK A COLO TOUR TOUR	0.350
	GOGGGCTGGTGTGATGCAGGCTGTGCGGGGGGGGGGGGGG	2400
15	PROCESTED BURGAN SAGEDEGGGGGGGGGTGATGGATGAGAT GAGAT GA	0450
	ATOSCOSCICAAOGGOCOGTOGTOGGTGGTT LICOGGTGATGAGGC LA A V N G P S S V V L S G D E A	2500
20	CGCCGTGCTGCAGGCCGCCGGGGAAGTGGACGCGGCTGGTGA A V L 2 A A E G L G K W T R L A	2550
	CCAGCCACGCGTTCCACTCCGCCCGTATGGAACCCATGCTGGAGGAGTTC T S R A F R S A R M E P M L E E F CGGGCGGTCGCCGAGGCCTGACCTACCGGACGCCGCAGGTCTCCATGGC	2600 2650
25	B A V A B B 1 T Y R T P Q V S M A CONTROL OF THE CO	
	V G D Q V T T A E Y W V R Q V R ACACGGTCCSCTTCGGCGAGCAGGTGGCCTCGTACGAGGACGCCGTGTTC	2750
	D T V F F G E Q V A S Y E D A V F GTCGAGCTGGCGAGGTGTCGA	2800
30	V E L G A D R S L A R L V D G V A GATGCTGCACGGGGACCACGAAATCCAGGCCGGGATGGGGGCCCTGGCCC	2850
	M L E G D E E I Q A A I G A L A ACCTGTATGTCAACGGCTCACGGTCGACTGGCCGCGCTCCTGGGCGAT	2900
35	H L Y U N G V T V D W P A L L G D GOTOOGGAAAAAGGGGTGGAGGACATAGGGCTTCCAGGACA	2950
	A P A T R V L D L P T Y A F Q H Q GOGCTACTGGCTCGGGCTCCCCGGGCCACTGGGGCCACC R Y W L E S A P P A T A D S G H	3000
40	COGTOCTOGGOACOGGAGTCGCCGGGTCGCCGGGGGGGGTGTTC	3050
	ACGGGTCCCGTGCCGCGCGGTGTTCATCGCCGAACT T G P V P A G A D R A V F I A E L	3100
	GGOGCTOGCOGCOGCACGCCACGGCCACGGTCGAACAGCTCG A L A A A B A T D C A T V E Q L	
45	ACGTCACCTCCGTGCCGGGGGGGGGGGGGGGGGGGGGGG	
	ACCTGGGTCGACGAAACCGGCGGCGGCGGCGGCGGCGGTCGACGTCGACGTCGACGGGGGGGG	
50	TRVGDAPWTLHAEGVL	3300
	BOCCCGGCCGCGTGCCCCAGCCCGAAGCCGTCGACACCGCCTGGCCCCCGCCCCCCCC	
55	FIGGREGOGGETGCCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	3450
er er	Q V F V E A E V D S P D G F V A	3500
	A PORT OF A TOP OF A	
60	I P T G W R D L A V H A S D A T V	

WO 00/20601 PCT/US99/22886

ontroppour protectiva estado 303A 9A 3T on et det de Aportoppour 2003 - Diva Air Dil Dil Bill Bill Dil Gillon Dil Vill Bill A 07700A0367600AAKKT3006F79CT0A003.33A6T03675A030T3.3659 3077.3h3.3ggn.goggg.ggggganggggan.ananana3232303h03h33.3g LEWENTABAHYCGACE 10 SKOGGONOGKACOCCACKARARANCOCACROCOCACROCACKARANING 18860 DEATOSTOCA DA CONCON CONSCINENTA GOSCO GOGOTOA COSCOCA DE 1989 15 TRTAQNEHPGRIHLIE BOADDADDDDDACACOCCACTOCCCCTCACCCAASTCACCACCCCCCCCACC 4050 20 HHPHTPLPLTQLTTLE ODDATOROGAOGRACIACACAGOAGCAGAACACCACORACACACACCACAR COTCARCOCCAACCACCCATCATCACCGGCGGCTCCGCACCCTCG 4200 25 INPNHALLITEGSGTL A G I L A R H L N H P H T Y L L S CGCACACCACCCCCACCACCACCCGGCACCCACATTCCCCTGCGACCT 4300 R T P P P T T P G T H I P C D L 30 CACCGACCCCAAATCACCCAAGCCCTCACCCACATACCACAACCCC 4350 T D P T Q I T Q A D T H T F Q P L T G J F H T A A T L D C A T L MACOTORODOCORRADOCORODADOROCACORODARA GODORA GODORO 4450 35 и втерявтить фекара GSCCTSGCMCCTCCACCACACCCAAAAACCAACCCTCACCCACTTCS 4500 A W H L H H H T Q N Q P L C H F 40 LYSSAAATLGSPGQAN Y A A A N A F L D A L A T H R H CCAMBGACAACCGGCACCACGATGGCCTGGGGCATGTGGCACACCACCA 4650 Q G Q P A T T I A W G M X H T T CCACACTCACCAGCCAACTCACCGACAGCGACCGCGACCGCATCCGCCGC 4700 45 TTLTSQLTDSDRIRR GGCGGCTTCCTGCCGATCTCGGACGACGAGGGCATGC G G F L P I S D D E G M

The *Nhel-Xhol* hybrid FK-506 PKS module 8 containing the AT domain of module 13 of rapamycin is shown below.

50

```
POST GRAMA A BOARD ROOM ON STROTT DIGGORACT SEED OF SACRACACT REC
FOR A FINE OF THE AFTER THE DISCUSSION AFEED OF
                GACGA CGA DETTOAA GGAA DTOĞĞOATICGA OTOGOTEA OOGOGG 3000
               5
        TAVESEPTERALAAPLS
          10
            - E la graa e vaa e taa
        AAABDEFLAIMGMA
        CONTRACTOR AND THE RESERVE OF THE STATE O
15
       I A L Y S P S P D A I S K T F V F
        CACGGGGGCTTCCTCGACGGTGGGACGGGCTTCGACGGGGGCTTCTTCGG
20
        B G G F L D G A T G F D A A F F G
        3ATQXGCCCCAAGCCCCTGCCCXTGGACCCGCAGCAXCCGCTCCC
       I B P K E A L A M D P Q Q P V 1
TGGAGAGGTCCTGGGAGGCGTTCGAAAGCGGGGGCATCAGGGGGAAGGGG 890
        L E T S W E A F E S A G I T F S A
25
       GCGCGGGGCAGCGACACCGGCGTGTTCATCGGCGCGTTCTCCTACGGGTA 850
        ARGSDTGVFIGAFSY31
       CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTCGCAGACCA 960
         B T G A D T N G F G A T G S Q
       GCGTGGTGTGCGGGCGGCGTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950
30
       S T L S G R L S Y F Y G L E G F S
       GTCAGGTGGACAGGGCTGCTCGTCGTCACTGGTCGCCCTGCACCAGGC 1000
        Y T Y D T A C S S S L Y A L H Q A
       AGGGCASTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTCGGCGGTG 1050
        35
          V T V M A S P G G F V E F S R Q B
       39997799999683426669966994A6693772389609<mark>99</mark>9666664883 1150
       G L A P E G R A K A F S A S A S S
       40
        T S F A E G A G A L V V E R I
      ACGOGGAGGOCACACGCCTCCTCGCCCTCGTACGCGGCTCCCCGG 1250
       DAERHGHTVLALVRGSA
       GCTAACTCCGACGGCGCGTCGAACGGTCTGTCGGCCCCGAACGGCCCCTC 1300
        A N S D G A S N G L S A P N G P S
45
       CCAGGAACGCGTCATCCACCAGGCCCTCGCGAACGCGAAACTCACCCCCG 1350
        QERVIHQALANAKLTE
      DECATOGA GUOGOA GEOGOTGOT DOCGACGTACGGA CAGGACCGGGGGAC - 1450
50
        PIEAÇALUATYGQDBA
       GCCCCTGCTCCTCGGCTCGCTGAAGTCGAACATCGGGCACGCCCAGGCCG 1500
         F t t t g s r k s N I G H A Q A
       COTORGOOTOGOOGGATCATCAMGRIGGIGGROGGCATCCGGCACGGG-18E0
       A S B T A G I I F M V Q A I R H S
       SAAKT BOORTOGA DADTGOA BAGGGGWOSA GOOGT OG BOGGA BOT GA BTG - 1 600
55
        ELFFTLHADEPSPHVDW
       PARTS SOURCE TO SETTING A GOT SOTTIGAR GET 06 GC 00 GC 00 TG 60 CG GG GA. 1650
        COGGTOGOCOGOGOGOGOTGCCGTCTCGTCGTTCGGGGGTGAGCGGCACG 1700
60
       T G R F R R A A V S S F G V S G T
```

	AACGCCCACATCATCCTTGAGGCAGGACGGGTCGAAAACGGGACGGGTCGA	. 1780 :
	- 937A33A3DGATGGA3GCAGGACCGGTCGAAGTAGGACCGGTCGAGGTG - A T A I E A G P M E M G E M E A	1300
5	- GROUGUTOCOOGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1850
	CTCGTGTCGGGGGGTTCCCCGGGAGGACTCGACGAGCAGATCGGGGGGCCT	
10	George CTATET COACACCEGO CO COGO COTOGACCEGO COGO COTOGO COGO COGO COGO COGO COGO	
	AGACACTGSCCGGSCGTAGGCACTTCACCCACGGGGCGTAGTGCTCGGG	
	SACACOSTOATOGGOOCTOOOCOGGOOGGOOGGOOGGOOGACGOOTOTT E T M 1 G A P F A D Q A D E L V F	
15	THE TRACEOUS TO A CARD CONTROL OF THE CONTROL OF TH	2100
	CODATTOGTOGGTGUTGTTCGCCGAGGGGGGGGGGGGGGGGGGGGGGGGG	
20	TTGCGCSAGTTCGTCGACTCGGATCTTCACCGGTTCTGGATGATCCGGC L R E F V D W D L F T V L D D P A	
	GSTGGTGGACCGGGTTGATGTGGTCCAGCCCGCTTCCTGGGCGATGATGG	
	TTTCCCTGGCCGGGTGTGGCAGGCGGGGGGGGGGGGGGG	2300
25	ATCGGCCATTCGCAGGGTGAGATCGCCGCAGCTTGTGTGGCGGGTGCGGT	
	GTCACTACGCGATGCCGCCGGATCGTGACCTTGCGCAGCCAGGCGATCG E L R D A A R I V T L R S Q A I	
30	CCCGGGGCCTGGCGGGCGGGGGGGGGGGATGGCATCCGTCGCCCTGCCCGCG A R G L A G R G A M A S V A L P A	
	CAGGATGTCGAGCGGGGCCTGGATCGCCGCCCACAACGGGCC Q D V E L V D G A W I A A H N G P	
	CGCCTCCACCGTGATCGCGGGGCACCCCGGGAAGCGGTCGACCATGTCCTCA A S T V I A G T P E A V D H V L	2550
35	COGCTCATGAGGCACAAGGGGTGCGGGGGGGGGGGGGGGG	
	SCCTCGCACACCCCGCGACCTCGACGTGATCCGCGACGAACTACTCGACAT A S H T P H V E L I R D E L L D I	2650
40	UNOTAGOGACAGOTOGOAGACOCOGOTOGTGCCGTGCCGTGC	
	TGGACGGCACCTGGGTCGACAGCCCGCTGGACGGGGAGTACTGGTACCGG V D G T W V D S P L D G E Y W Y R	
	AACCTGCGTGAACCGGTCGGTTTCCACCCCGCCGTCAGCCAGTTGCAGGC N L R E P V G F H P A V S Q L Q A	2800
45	CCAGGGCGACACCGTGTTCGTCGAGGTCAGCGCCAGCCCGGTGTTGTTGC Q G D T V F V E V S A S P V L L	2850
	AGGOGATGGACGATGTCGTCACGGTTGCCACGCTGCGTCGTGACGAC	2900
50	GGCGACGCCACCCGGATGCTCACCGCCCTGGCACAGGCCTATGTCCACGG G D A T R M L T A L A Q A Y V H G	2950
	COTCACOGTOGACTGGCCCGCCATCCTCGGCACCACCACCACACCCGGGTAC	3000
	TGGACCTTCCGACCTACGCCTTCCAACACCAGCGGTACTGGCTCGAGTCG L	3050
55	GOTOCCCGGCCACGGCCGACTCGGGCACCCGGAGT A P F A T A D S G H P V L G T G V	3100
	TROCONTOS COSSONOS COS COSONOS TO THE TROCOS COSONOS COSONOS CONTRACTOR TO THE TOTAL COSONOS CONTRACTOR C	3150
60	GTGGGGACCGGGGGGTGTTCATCGCCGAACTGGGGGCTCGCCGCCGCCGAC	3200

		3250
	000247000000000000000000000000000000000	3355
		2232
5		
.)	- 009009ACGGGGGGGGCCTTCACCCCCACACCGGGGCCCCGGGACCGCGA	3350
	AADGARRETVETRVGDA	
	- COGTGGACGCTGCACGCCGAGGGGGTTCTGCGCCCGCGCGCG	3400
	PWTLHAEGVLRPGRVPD	3.00
	-	
1./	3000GAAGOCGTC3ACACGCCT6600CCCGCCGGGGGGGGGGTG000GG	3450
10	PEAVDTAWPPPGAVPA	
	ACGGGCTGCCCGGGGGGTGGCGACGCGGGGCCAGGTCTTCGTGGAAGGC	3500
	D G L P G A W R R A D Q V F V E A	5 5 6 6
	GAAGTOGACAGCCCTGACGGCTTCGTGGCACACCCCGACCTGCTCGACGC	3550
	LEV D S P D G F V A H P D L L D A	
15	SCECETAEDO A CODO A DODO A DEBO CA SOCIO A SOCIO A SOCIO A TODO CONTRO COME	3600
		2 2 5 0
	ACCTOGOGGTGCACGCGTCGGACGCCACCGTGCTGCGCGCCTGCCTCACC	3650
	DLAVHASDATVLRACLT	
	CGCCCCGACACTGCTCGTGGAGCTCGCCGCCTTCGACGGTGCCGGAAT	3700
20		3,00
20	R R D S G V V E L A A F D G A G M	
	GCCGGTGCTCACCGCGAGTCGGTGACGCTGGGCGAGGTCGCGTCGGCAG	3750
	PVLTAESVTLGEVASA	
	GCGSATCCGACGACTCGGACGCTCTGCTTCGGCTTGAGTGGTT3CCGGTG	3800
		2000
	F G S D E S D G L L R L E W L F V	
25	GCGGAGGCCCACTACGACGGTGCCGACGAGCTGCCCGAGGGCTACACCCT	3850
	A E A H Y D G A D E L P E G Y T L	
	CATCACCGCCACACACCCCGACGACCCCGACGACCCCACCAACCCCCACA	3900
		3900
	TTATHPODPDDPTNPH	
	ACACACCCACACGCACCCCACACACACACACGCGTCCTCACCGCCCTC	3950
30	NTPTRTHTOTTRVLTAL	
20		4000
	CAACACCACCTCATCACCACCACCACCACCCTCATCGTCCACACCACCAC	4000
	QHHLITTNHTLIVHTTT	
	CGACCCCCCAGGCGCCGCGCTCACCGGCCTCACCCGCACCGCACAAAACG	4050
	DPPGAAVTGLTRTAQN	
35		
7 3	130000000000000000000000000000000000000	1100
20	ARCACCCCGGCCGCATCCACCTCATCGAAACCCACCACCCCCCACACCCCCA	4100
20	AACACCCGGCCGCATCCACCTCATCGAAACCCACCACCCCCACACCCCA E H P G R I H L I E T H H P H T P	4100
20	EHPGRIHLIETHHPHTP	
20	E H P G R I H L I E T H H P H T P TTOCCCCTCACCCAACTCACCCACCACCACCAACCCAAC	4100 4150
	E H P G R I H L I E T H H P H T P CTCCCCCTCACCAACCACCACCAACCCACCAACCCACCAACCCACCA	4150
	E H P G R I H L I E T H H P H T P TTOCCCCTCACCCAACTCACCCACCACCACCAACCCAAC	
40	E H P G R I H L I E T H H P H T P CTCCCCCTCACCAACCACCACCAACCCACCAACCCACCAACCCACCA	4150
	E H P G R I H L I E T H H P H T P CTCCCCCTCACCAACCCACCACCCACCCACCCACCCAC	4150 4200
	E H P G R I H L I E T H H P H T P CTCCCCCTCACCAACTCACCACCTCCACCAACCCACCTACGCCTCAC L P L T Q L T T L H Q P H L F L T CAACAACACCCTCCACACCCCCACCTCACCCCCATCACCAC	4150 4200
	E H P G R I H L I E T H H P H T P CTCCCCCTCACCAACTCACCACCTCCACCAACCCACCTACGCCTCAC L P L T Q L T T L H Q P H L F L T CAACAACACCCTCCACACCCCCACCCCCATCACCACCACC	4150 4200 4250
	E H P G R I H L I E T H H P H T P CTCCCCCTCACCAACTCACCACCTCCACCAACCCACCTACGCCTCAC L P L T Q L T T L H Q P H L F L T CAACAACACCCTCCACACCCCCACCTCACCCCCATCACCAC	4150 4200 4250
	E H P G R I H L I E T H H P H T P CTCCCCCTCACCAACTCACCACCTCCACCACCCCACC	4150 4200 4250
40	E H P G R I H L I E T H H P H T P CTCCCCCTCACCAACTCACCACCTCCACCACCCACCCA	4150 4200 4250 4300
	E H P G R I H L I E T H H P H T P CTCCCCCTCACCAACTCACCACCTCCACCACCCCCCCC	4150 4200 4250 4300
40	E H P G R I H L I E T H H P H T P CTCCCCCTCACCAACTCACCACCTCCACCACCCCACC	4150 4230 4250 4300 4350
40	E H P G R I H L I E T H H P H T P CTCCCCCTCACCAACTCACCACCTCCACCACCCCACC	4150 4230 4250 4300 4350
40	E H P G R I H L I E T H H P H T P TTOCCCCTCACCAACTCACCACCTCCACCACCCCCCCCCC	4150 4230 4250 4300 4350
40	E H P G R I H L I E T H H P H T P TTOCCCCTCACCAACTCACCACCTCACCAACCCAACCC	4150 4230 4250 4300 4350 4400
40 45	E H P G R I H L I E T H H P H T P TTOCCCCTCACCAACTCACCACCTCACCAACCCTACCCT	4150 4230 4250 4300 4350 4400
40	E H P G R I H L I E T H H P H T P TTOCCCCTCACCAACTCACCACCTCACCAACCCTACCCT	4150 4230 4250 4300 4350 4400
40 45	E H P G R I H L I E T H H P H T P TTOCCCCTCACCAACTCACCACCTCACCAACCCTACCCT	4150 4230 4250 4300 4350 4400
40 45	E H P G R I H L I E T H H P H T P TTOCCCCTCACCAACTCACCACCTCACCAACCCTACCCT	4150 4230 4250 4300 4350 4400
40 45	E H P G R I H L I E T H H P H T P TTCCCCCTCACCAACTCACCACCTCCACCAACCCTACGCTCAC L P L T Q L T T L H Q P H L F L T CAACAACACCCTCCACCAACCCCCAACCCCACCCCA	4150 4230 4250 4300 4350 4400 4500
40 45	E H P G R I H L I E T H H P H T P TTCCCCCTCACCCACTCACCACCTCCACCAACCCTACGCTCAC L P L T Q L T T L H Q P H L F L T CAACAACACCCTCCACCACCCCCACCCCCACCCCCACCCCCACCCCCACCCC	4150 4230 4250 4300 4350 4400
40 45 50	E H P G R I H L I E T H H P H T P TTCCCCCTCACCCACTCACCACCTCCACCACCCTACGCTCAC L P L T Q L T T L H Q P H L F L T CAACAACACCCTCCACCACCCCCACCCCACCCCACC	4150 4230 4250 4300 4350 4400 4500
40 45	E H P G R I H L I E T H H P H T P TTCCCCCTCACCCACTCACCACCTCCACCACCCTACGCTCAC L P L T Q L T T L H Q P H L F L T CAACAACACCCTCCACCACCCCCACCCCACCCCACC	4150 4230 4250 4300 4350 4400 4500
40 45 50	E H P G R I H L I E T H H P H T P TTOCCCCTCACCCACTCACCACCTCCACCAACCCTACGCTCAC L P L T Q L T T L H Q P H L F L T CAACAACACCCTCCACCACCCCCACCCCCACCCCCACCCCCACCCCCACCCC	4150 4230 4250 4300 4350 4400 4500 4550
40 45 50	E H P G R I H L I E T H H P H T P TTOCCCCTCACCCACTCACCACCTCACCAACCCTACGCTTCAC L P L T Q L T T L H Q P H L F L T CAACAACACCCTCCACCACCCCCACCCCCACCCCCACCCCCACCCCCACCCC	4150 4230 4250 4300 4350 4400 4500 4600
40 45 50	E H P G R I H L I E T H H P H T P TTOCCCCTCACCCACTCACCACCTCCACCACCCTACGCTCAC L P L T Q L T T L H Q P H L F L T CAACAACACCCTCCACCACCCCCACCCCACCCCCACCCCCACCCCCACCCCC	4150 4230 4250 4300 4350 4400 4500 4550
40 45 50	E H P G R I H L I E T H H P H T P TTOCCCCTCACCCACTCACCACCTCACCAACCCTACGCTTCAC L P L T Q L T T L H Q P H L F L T CAACAACACCCTCCACCACCCCCACCCCCACCCCCACCCCCACCCCCACCCC	4150 4230 4250 4300 4350 4400 4500 4600
40 45 50	E H P G R I H L I E T H H P H T P TTOCCCCTCACCCACTCACCACCTCCACCAACCCTACGCTCAC L P L T Q L T T L H Q P H L P L T CAACAACACCCTCCACCACCCCCACCCCCATCACCCCCACCCCCACCCCCACCCCCACCCCCACCCCCC	4150 4230 4250 4300 4350 4400 4500 4600 4650
40 45 50	E H P G R I H L I E T H H P H T P TTOCCCCTCACCCACTCACCACCTCCACCACCCTACGCTCAC L P L T Q L T T L H Q P H L F L T CAACAACACCCTCCACCACCCCCACCCCACCCCCACCCCCACCCCCACCCCC	4150 4230 4250 4300 4350 4400 4500 4600 4650

15

20

2.5

30

35

P D D B G M

Example 3

PCT/US99/22886

Recombinant PKS Genes for 13-desmethoxy FK-506 and FK-520

The present invention provides a variety of recombinant PKS genes in addition to those described in Examples 1 and 2 for producing 13-desmethoxy FK-506 and FK-520 compounds. This Example provides the construction protocols for recombinant FK-520 and FK-506 (from Streptomyces sp. MA6858 (ATCC 55098), described in U.S. Patent Nos. 5.116,756, incorporated herein by reference) PKS genes in which the module 8 AT coding sequences have been replaced by either the rapAT3 (the AT domain from module 3 of the rapamycin PKS), rapAT12, eryAT1 (the AT domain from module 1 of the erythromycin (DEBS) PKS), or ervAT2 coding sequences. Each of these constructs provides a PKS that produces the 13-desmethoxy-13-methyl derivative, except for the rapAT12 replacement, which provides the 13-desmethoxy derivative, i.e., it has a hydrogen where the other derivatives have methyl.

Figure 7 shows the process used to generate the AT replacement constructs. First, a fragment of ~4.5 kb containing module 8 coding sequences from the FK-520 cluster of ATCC 14891 was cloned using the convenient restriction sites SacI and SphI (Step A in Figure 7). The choice of restriction sites used to clone a 4.0 - 4.5 kb fragment comprising module 8 coding sequences from other FK-520 or FK-506 clusters can be different depending on the DNA sequence, but the overall scheme is identical. The unique SacI and SphI restriction sites at the ends of the FK-520 module 8 fragment were then changed to unique Bgl II and NsiI sites by ligation to synthetic linkers (described in the preceding Examples, see Step B of Figure 7). Fragments containing sequences 5' and 3' of the AT8 sequences were then amplified using primers, described above, that introduced either an AvrII site or an NheI site at two different KS/AT boundaries and an XhoI site at the AT/DH boundary (Step C of Figure 7). Heterologous AT domains from the rapamycin and erythromycin gene clusters were amplified using primers, as described above, that introduced the same sites as just described (Step D of Figure 7). The fragments were ligated to give hybrid modules with in-frame fusions at the KS/AT and AT/DH boundaries (Step E of Figure 7). Finally, these hybrid modules were ligated into the BamHI and PstI sites of the WO 00/20601

10

KC515 vector. The resulting recombinant phage were used to transform the FK-506 and FK-520 producer strains to yield the desired recombinant cells, as described in the preceding Examples.

The following table shows the location and sequences surrounding the engineered site of each of the heterologous AT domains employed. The FK-506 hybrid construct was used as a control for the FK-520 recombinant cells produced, and a similar FK-520 hybrid construct was used as a control for the FK-506 recombinant cells.

Hecologous AT	Enzyme	Location of Engineered Site
FK-506 AT8	AvrII	GGCCGT@cgcgcCGTCTCGTCGTTC
(hydroxymalonyl)	1	GRPRRAAVSSF
(nydroxymaionyi)	Nheĭ	ACCONSCATOCCCCCATAGGTGAGCGgatogeC
	ivnei	TQHPAMGERLA
		TACGCCTTCCAGCGGCGGGCCTACTGGatcgag
	Xhol	YAFQARPYWIE
rapamycin AT3	AvrII	GACCGG <u>becest</u> DGGGCGGGCGTGTCGTTC
(methylmalonyl)		DRPRRAGVSSF
	Nhel	TGGCAGTGGCTGGGGATGGGCAGTGCactacaG
		W Q W L G M G S A L R
	XhoI	TACGCCTTCCAACACCAGCGGTACTGGgtcgag
		Y A F Q H Q R Y W V E
rapamycin AT12	AvrII	GGCCGAacaccGGGCAGGCGTGTCGTCCTTC
(malonyl)		GRARRAGVSSF
	NheI	TCGCAGCGTGCTGGCATGGGTGAGGAactggcC
		S Q R A S M G E E L A TACGCCTTCCAGCACCAGCGCTACTGGctcgag
	XhoI	Y A F O H C R Y W L E
DEBS AT1	AvrII	SCECEA DE DE CONTROL DE LA C
	AWTII	A R P R P A G V S S F
(methylmalonyl)		TGGCAGTGGGCGGSCATGGCCGTCGAcctactC
	NheI	WOWAGMAVDLL
		TACCCGTTCCAGCGCGAGCGCGTCTGGctcgaa
	XhoI	Y P F Q R E R V W L E
DEBS AT2	AvrII	GACGGGgtgcgcGGGCAGGTGTGTCGGCGTTC
(methylmalonyl)		D G V R R A G V S A F
(memyimatonyi)	NheI	GCCCAGTGGGAAGGCATGGCGCGGGAgttgttG
	Ivnei	AQWEGMARELL
		TATCCTTTCCAGGGCAAGCGGTTCTGG <u>ctgctg</u>
	XhoI	Y P F Q G K R F W L L

The sequences shown below provide the location of the KS/AT boundaries chosen in the FK-520 module 8 coding sequences. Regions where AvrII and NheI sites were engineered are indicated by lower case and underlining.

25

00ag0ag0ggrgak.htggrak.artggggggggggggggggakakeggkGgggagagg ASAVELLTSASPWESTCSFR A A T D E P P T T D A B T T L E A ADDAGARDETES AT DESPESSET DA BESSES PAR ADBAGOT CESTES ARBORROT DE STESSA A PAR ABBAGOT A STESSA A CANDAS A CA A F S F E A L D E Q I R R L R A Y L D T C P D Y D R Y A V A Q T L A R A T H F A H R A V L L G D T V I T T P P A C R P D AACTOGTOTTOCTCTACTCCGGCCAGGGCACCAGGATCCCGGATTGGGCCAAGCTACC E L V F V Y S G Q G C Q H P A N G E C L goscoccoatecestattogecgacgcotscoll grade and considerations and considerations and considerations and considerations and considerations and considerations are considerated as a consideration of the consideration and considerations are considerated as a consideration of the consideration and considerations are considerated as a consideration of the consideration of t TOOGOOGCOTTGACAACC 15 A A A H P V F A D A W H E A L R P L D W

The sequences shown below provide the location of the AT/DH boundary chosen in the FK-520 module 8 coding sequences. The region where an *MoI* site was engineered is indicated by lower case and underlining.

THE THE GRADIES TO A GRADIES AND A DESCRIPTION OF THE GRADIES AND A DESCRIPTION OF THE PROPERTY OF THE PROPER

The sequences shown below provide the location of the KS/AT boundaries chosen in the FK-506 module 8 coding sequences. Regions where AvrII and NheI sites were engineered are indicated by lower case and underlining.

TCGGCCAGGCGTGGCCGCGGACCGGCCGTgcgggcCGTGCGGCGGCGTCTCGTCGTTCGGG 30 S A R P W P R T G R P R R A A V S S F G GTGAGCGGCACCAACGCCCACATCATCCTGGAGGCCGGACCGGACCAGGAGGAGCCGTCG V S G T N A H I I L E A G P D Q E E P S GOAGAACCGGCCGGTGACCTCCCGCTGCTCGTGTCGGCACGGTCCCCGGAGGCACTGGAC A E F A G D L P L L V S A R S P E A L D GAGCAGATCGGGGGGCTGGGGGACTATCTCGACGCCGCCCCGGGGTGGACCTGGCGGCC 35 EQIGRLEDY LOAA PSVOLAA GTGGCGCGGACACTGGCCACGCGTACGCACTTCTCCCACCGCGCCGTACTGCTCGGTGAC V A R T L A T R T H F S H R A V L L G D ACCGTCATCACCGCTCCCCCCGTGGAACAGCCGGGCGAGCTCGTCTTCGTCTACTCGGGA TVITAPPVEQPGELVFVYSG CAGGGCACCCAGCATCCCGCGATGGGTGAGCGGCEcacCGCAGCCTTCCCCGTGTTCGCC Q G T Q H P A M G E R L A A A F P V F A GACCOGGACGTACCCGCCTACGCCTTCCAGCGGCGCCCTACTGGATCGAGTCCGCGCCG D P D V P A Y A F Q R R P Y W I E S A P 45

The sequences shown below provide the location of the AT/DH boundary chosen in the FK-506 module 8 coding sequences. The region where an *XhoI* site was engineered is indicated by lower case and underlining.

Example 4

10

Replacement of Methoxyl with Hydrogen or Methyl at C-15 of FK-506 and FK-520

The methods and reagents of the present invention also provide novel FK-506 and FK-520 derivatives in which the methoxy group at C-15 is replaced by a hydrogen or methyl. These derivatives are produced in recombinant host cells of the invention that express recombinant PKS enzymes the produce the derivatives. These recombinant PKS enzymes are prepared in accordance with the methodology of Examples 1 and 2, with the exception that AT domain of module 7, instead of module 8, is replaced. Moreover, the present invention provides recombinant PKS enzymes in which the AT domains of both modules 7 and 8 have been enanged. The table below summarizes the various compounds provided by the present invention.

	Compound	C-13	C-15	Derivative Provided
	FK-506	hydrogen	hydrogen	13, 15-didesmethoxy-FK-506
	FK-506	hydrogen	methoxy	13-desmethoxy-FK-506
15	FK-506	hydrogen	methyl	13,15-didesmethoxy-15-methyl-FK-506
	FK-506	methoxy	hydrogen	15-desmethoxy-FK-506
	FK-506	methoxy	methoxy	Original Compound FK-506
	FK-506	methoxy	methyl	15-desmethoxy-15-methyl-FK-506
	FK-506	methyl	hydrogen	13,15-didesmethoxy-13-methyl-FK-506
20	FK-506	methyl	methoxy	13-desmethoxy-13-methyl-FK-506
	FK-506	methyl	methyl	13,15-didesmethoxy-13,15-dimethyl-FK-506
	FK-520	hydrogen	hydrogen	13, 15-didesmethoxy FK-520
	FK-520	hydrogen	methoxy	13-desmethoxy FK-520
	FK-520	hydrogen	methyl	13,15-didesmethoxy-15-methyl-FK-520
25	FK-520	methoxy	hydrogen	15-desmethoxy-FK-520
	FK-520	methoxy	methoxy	Original Compound FK-520
	FK-520	methoxy	methyl	15-desmethoxy-15-methyl-FK-520
	FK-520	methyl	hydrogen	13,15-didesmethoxy-13-methyl-FK-520
	FK-520	methyl	methoxy	13-desmethoxy-13-methyl-FK-520
30	FK-520	methyl	methyl	13,15-didesmethoxy-13,15-dimethyl-FK-520

Example 5

Replacement of Methoxyl with Ethyl at C-13 and/or C-15 of FK-506 and FK-520

30

The present invention also provides novel FK-506 and FK-520 derivative compounds in which the methoxy groups at either or both the C-13 and C-15 positions are instead ethyl groups. These compounds are produced by novel PKS enzymes of the invention in which the AT domains of modules 8 and or 7 are converted to ethylmalonyl specific AT domains by modification of the PKS gene that encodes the module. Ethylmalonyl specific AT domain coding sequences can be obtained from, for example, the FK-520 PKS genes, the niddamycin PKS genes, and the tylosin PKS genes. The novel PKS genes of the invention include not only those in which either or both of the AT domains of modules 7 and 8 have been converted to ethylmalonyl specific AT domain but also those in which one of the modules is converted to an ethylmalonyl specific AT domain and the other is converted to a malonyl specific or a methylmalonyl specific AT domain.

Example 6

Neurotrophic Compounds

15 The compounds described in Examples 1 - 4, inclusive have immunosuppressant activity and can be employed as immunosuppressants in a manner and in formulations similar to those employed for FK-506. The compounds of the invention are generally effective for the prevention of organ rejection in patients receiving organ transplants and in particular can be used for immunosuppression following orthotopic liver transplantation.

20 These compounds also have pharmacokinetic properties and metabolism that are more advantageous for certain applications relative to those of FK-506 or FK-520. These compounds are also neurotrophic; however, for use as neurotrophins, it is desirable to modify the compounds to diminish or abolish their immunosuppressant activity. This can be readily accomplished by hydroxylating the compounds at the C-18 position using

25 established chemical methodology or novel FK-520 PKS genes provided by the present invention.

Thus, in one aspect, the present invention provides a method for stimulating nerve growth that comprises administering a therapeutically effective dose of 18-hydroxy-FK-520. In another embodiment, the compound administered is a C-18,20-dihydroxy-FK-520 derivative. In another embodiment, the compound administered is a C-13-desmethoxy and/or C-15-desmethoxy 18-hydroxy-FK-520 derivative. In another embodiment, the compound administered is a C-13-desmethoxy and/or C-15-desmethoxy 18,20-dihydroxy-FK-520 derivative. In other embodiments, the compounds are the corresponding analogs of

10

15

20

25

30

FK-506. The 18-hydroxy compounds of the invention can be prepared chemically, as described in U.S. Patent No. 5,189,042, incorporated herein by reference, or by fermentation of a recombinant host cell provided by the present invention that expresses a recombinant PKS in which the module 5 DH domain has been deleted or rendered non-functional.

The chemical methodology is as follows. A compound of the invention (~200 mg) is dissolved in 3 mL of dry methylene chloride and added to 45 µL of 2,6-lutidine, and the mixture stirred at room temperature. After 10 minutes, tert-butyldimethylsilyl trifluoromethanesulfonate (64 µL) is added by syringe. After 15 minutes, the reaction mixture is diluted with ethyl acetate, washed with saturated bicarbonate, washed with brine. and the organic phase dried over magnesium sulfate. Removal of solvent in vacuo and flash chromatography on silica gel (ethyl acetate: hexane (1:2) plus 1% methanol) gives the protected compound, which is dissolved in 95% ethanol (2.2 mL) and to which is added 53 μL of pyridine, followed by selenium dioxide (58 mg). The flask is fitted with a water condenser and heated to 70°C on a mantle. After 20 hours, the mixture is cooled to room temperature, filtered through diatomaceous earth, and the filtrate poured into a saturated sodium bicarbonate solution. This is extracted with ethyl acetate, and the organic phase is washed with brine and dried over magnesium sulfate. The solution is concentrated and purified by flash chromatography on silica gel (ethyl acetate: hexane (1:2) plus 1% methanol) to give the protected 18-hydroxy compound. This compound is dissolved in acetonitrile and treated with aqueous HF to remove the protecting groups. After dilution with ethyl acetate, the mixture is washed with saturated bicarbonate and brine, dried over magnesium sulfate, filtered, and evaporated to yield the 18-hydroxy compound. Thus, the present invention provides the C-18-hydroxyl derivatives of the compounds described in Examples 1 - 4.

Those of skill in the art will recognize that other suitable chemical procedures can be used to prepare the novel 18-hydroxy compounds of the invention. See, e.g., Kawai et al., Jan. 1993, Structure-activity profiles of macrolactam immunosuppressant FK-506 analogues, FEBS Letters 316(2): 107-113, incorporated herein by reference. These methods can be used to prepare both the C18-[S]-OH and C18-[R]-OH enantiomers, with the R enantiomer showing a somewhat lower IC₅₀, which may be preferred in some applications. See Kawai et al., supra. Another preferred protocol is described in Umbreit and Sharpless, 1977, JACS 99(16): 1526-28, although it may be preferable to use 30 equivalents each of

SeO₂ and t-BuOOH rather than the 0.02 and 3-4 equivalents, respectively, described in that reference.

All scientific and patent publications referenced herein are hereby incorporated by reference. The invention having now been described by way of written description and example, those of skill in the art will recognize that the invention can be practiced in a variety of embodiments, that the foregoing description and example is for purposes of illustration and not limitation of the following claims.

Claims

1. An isolated nucleic acid that encodes a CoA ligase, a non-ribosomal peptide synthesize, or a domain of an extender module of a polyketide synthase enzyme that synthesizes FK-520.

5

2. The isolated nucleic acid of claim 1 that encodes an extender module, said module comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.

10

3. The isolated nucleic acid of claim 1 that encodes an open reading frame, said open reading frame comprising coding sequences for two or more extender modules, each extender module comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.

15

4. The isolated nucleic acid of claim 1 that encodes a gene cluster, said gene cluster comprising two or more open reading frames, each of said open reading frames comprising coding sequences for two or more extender modules, each of said extender modules comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.

20

5. The isolated nucleic acid of claim 2, wherein at least one of said domains is a domain of a module of a non-FK-520 polyketide synthase.

25

- 6. The isolated nucleic acid of claim 1, wherein said nucleic acid is a recombinant vector capable of replication in or integration into the chromosome of a host cell.
 - 7. The isolated nucleic acid of claim 6 that is selected from the group consisting of cosmid pKOS034-120, cosmid pKOS034-124, cosmid pKOS065-M27, and cosmid pKOS065-M21.

30

8. The isolated nucleic acid of claim 5, wherein said non-FK-520 polyketide synthase is rapamycin polyketide synthase. FK-506 polyketide synthase, or erythromcyin polyketide synthase.

15

- 9. A method of preparing a polyketide, said method comprising transforming a host cell with a recombinant DNA vector of claim 6, and culturing said host cell under conditions such that said polyketide synthase is produced and catalyzes synthesis of said polyketide.
 - 10. The method of claim 9, wherein said host cell is a Streptomyces host cell.
- 11. The method of claim 9, wherein said polyketide is selected from the group consisting of FK-520, 13-desmethoxy-FK-520, and 13-desmethoxy-FK-506.
 - 12. A recombinant host cell that expresses a recombinant polyketide synthase selected from the group consisting of: (i) an FK-520 polyketide synthase in which at least one AT domain is replaced by an AT domain of a non-FK-520 polyketide synthase; (ii) an FK-506 polyketide synthase in which at least one AT domain is replaced by an AT domain of a non-FK-506 polyketide synthase; (iii) an FK-520 polyketide synthase in which at least one DH domain has been deleted; (iv) an FK-506 polyketide synthase in which at least one DH domain has been deleted.
- 20 13. The recombinant host cell of claim 12 that expresses an FK-520 polyketide synthase in which an AT domain of module 8 has been replaced by an AT domain that binds malonyl CoA, methylmalonyl CoA, or ethylmalonyl CoA.
- 14. The recombinant host cell of claim 12 that expresses an FK-506 polyketide synthase in which an AT domain of module 8 has been replaced by an AT domain that binds malonyl CoA, methylmalonyl CoA, or ethylmalonyl CoA.
 - 15. The recombinant host cell of claim 13, wherein a DH domain of module 5 or module 6 has been deleted.
 - 16. The recombinant host cell of claim 14, wherein a DH domain of module 5 or module 6 has been deleted.

30

17. A recombinant host cell that comprises recombinant genes coding for enzymes sufficient for synthesis of ethylmalonyl CoA or 2-hydroxymalonyl CoA.

18. A polyketide having the structure

5

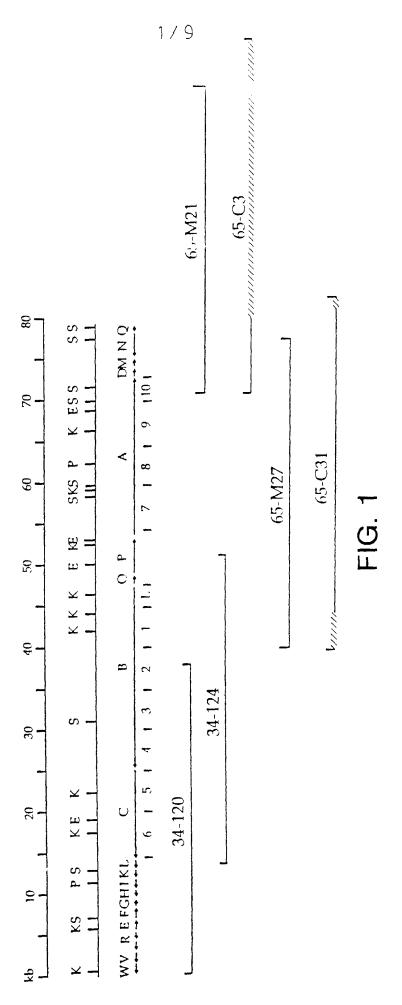
wherein, R₁ is hydrogen, methyl, ethyl, or allyl; R₂ is hydrogen or hydroxyl, provided that when R₂ is hydrogen, there is a double bond between C-20 and C-19; R₃ is hydrogen or hydroxyl; R₄ is methoxyl, hydrogen, methyl, or ethyl; and R₅ is methoxyl, hydrogen, methyl, or ethyl; but not including FK-506, FK-520, 18-hydroxy-FK-520, and 18-hydroxy-FK-506.

19. The polyketide of claim 18 that is 13-desmethoxy-FK-506.

15

10

20. The polyketide of claim 18 that is 13-desmethoxy-18-hydroxy-FK-520.



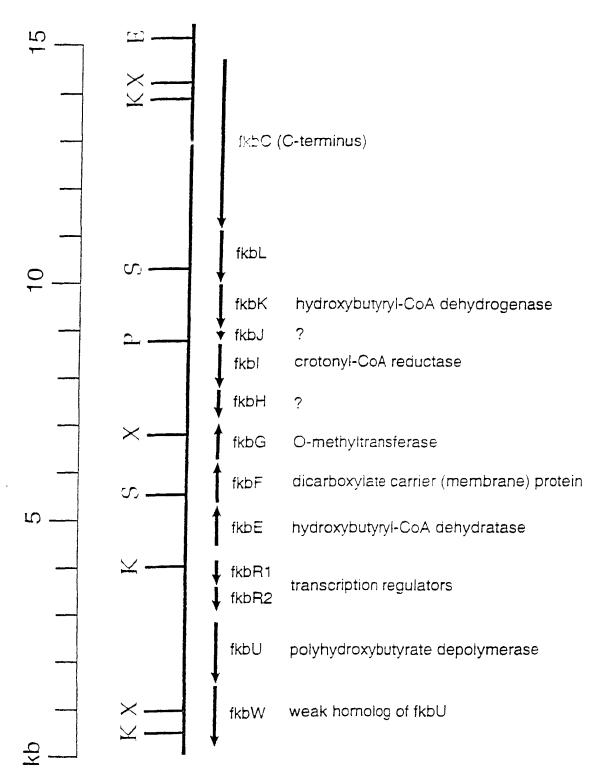


FIG. 3

FIG. 4

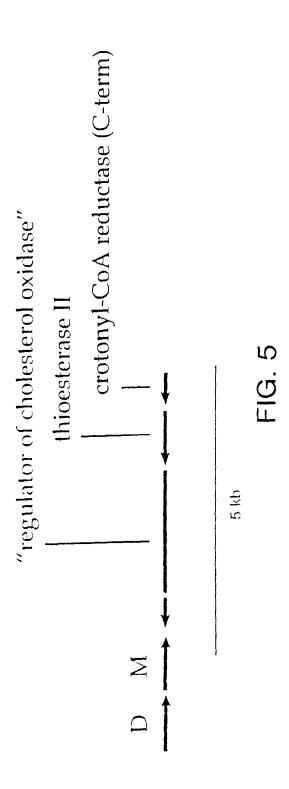
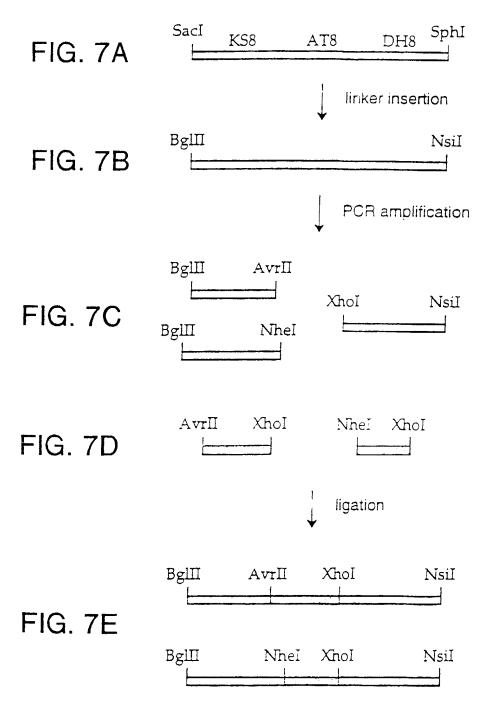


FIG. 6



INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism re on page 22 , line 31	eferred to in the description -33
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution American Type Cul	ture Collection
Address of depositary institution (including per and country)	y)
10801 University F Manassas, VA 2211 USA	
Date of deposit	Accession Number
20 September 1999	PTA-727, PTA-728 and PTA-729
C. ADDITIONAL INDICATIONS (leave blank if not applicate	ole) This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDICATION All designated States	ONS ARE MADE (if the indications are not for all designated States)
E. SEPARATE FURNISHING OF INDICATIONS (leav	e blank if not applicable)
The indications listed below will be submitted to the International Number of Deposit*)	Buteau later (specify the general nature of the indications e.g., "Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer	Authorized officer
Form PCT/RO/134 (July 1992)	

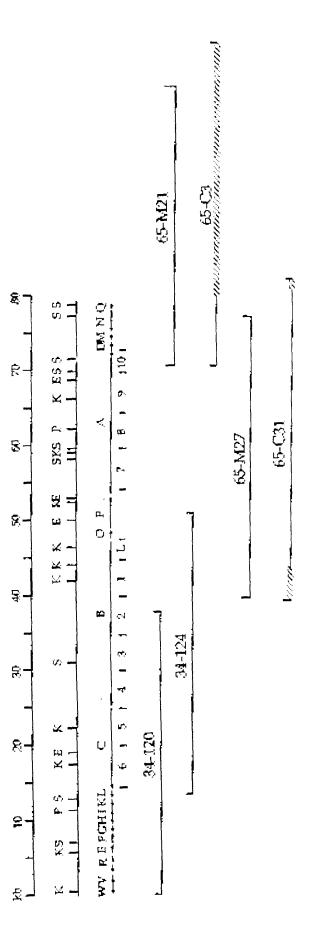
INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism re on page 22, line 3	ferred to in the description 1-33
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet $[X]$
Name of depositary institution American Type Culture Address of depositary institution (including postal code and country 10801 University Blvd Manassas, VA 22110-220)
USA Date of deposit 20 September 1999	Accession Number PTA-726
C. ADDITIONAL INDICATIONS (leave blank if not applicable	le) This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDICATION All designated States	NS ARE MADE (if the indications are not for all designated States)
E. SEPARATE FURNISHING OF INDICATIONS (leave	
The indications listed below will be submitted to the International Number of Deposit*)	Buteau later (specify the general nature of the indications e.g., "Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

Form PCT/RO/134 (July 1992)





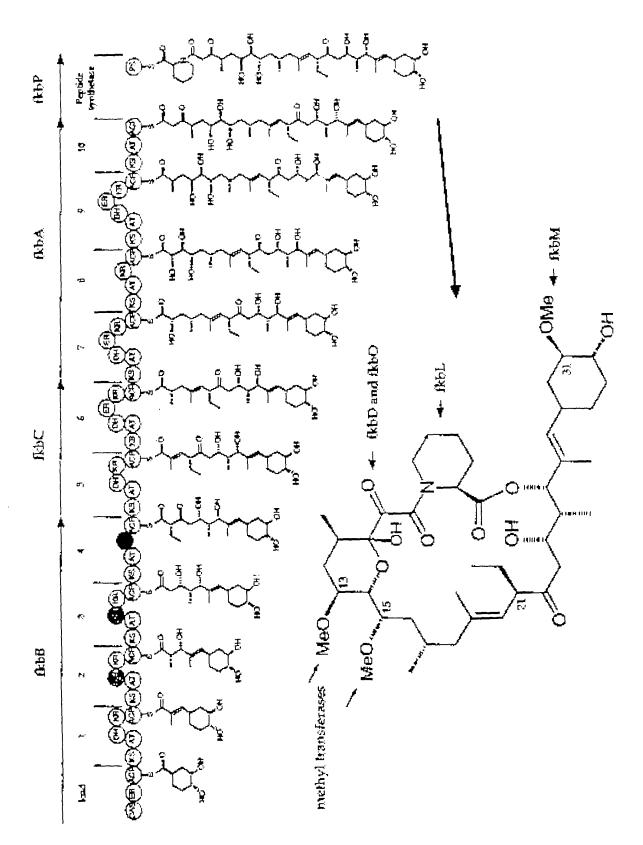


Figure 2

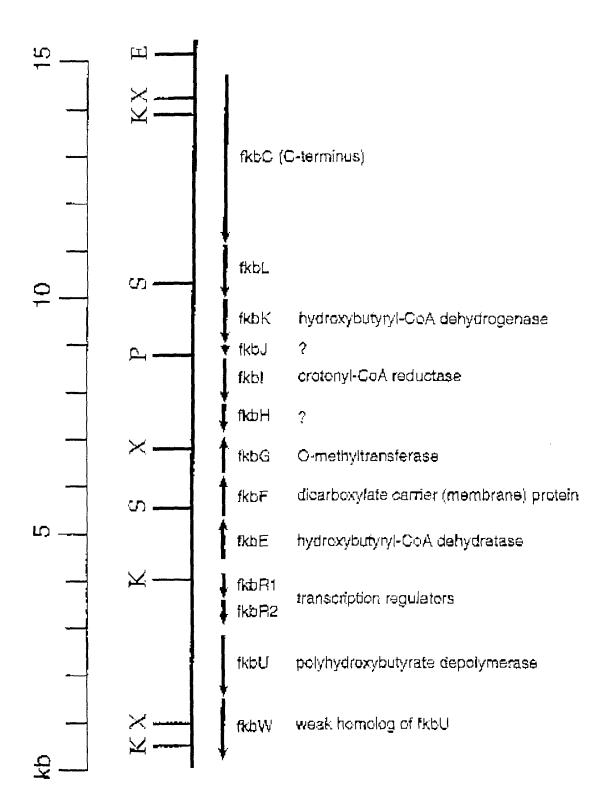


Figure 3

Figure 4

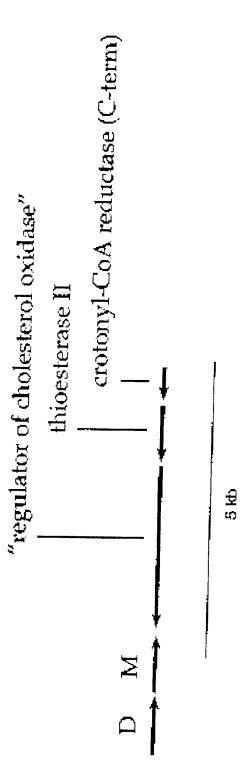


Figure 5

Figure 6

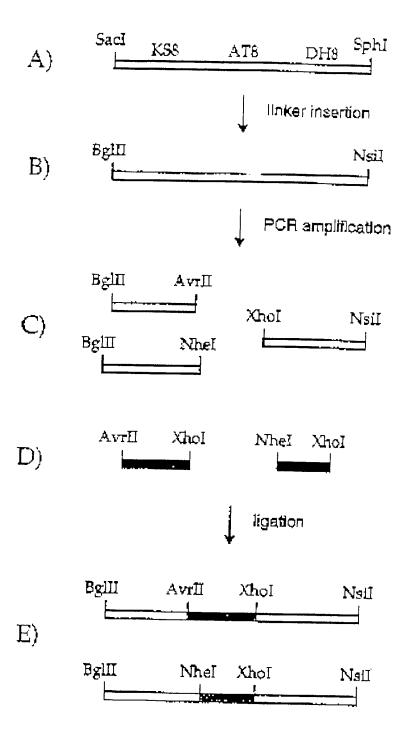


Figure 7

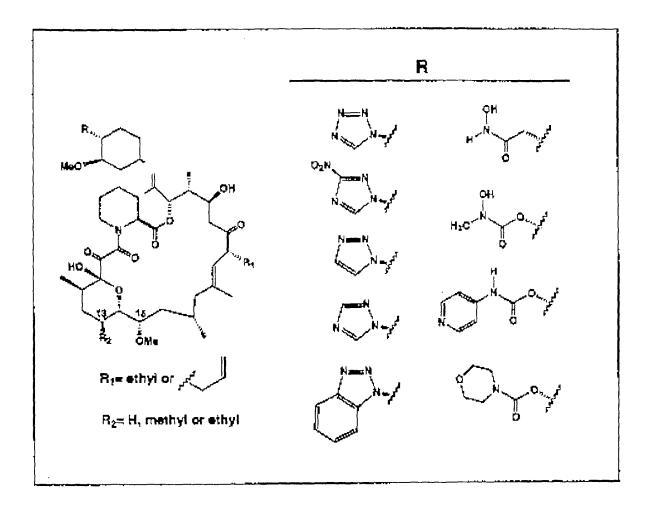


Figure 8 Part A

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7:

C12N 15/52, 15/54, 15/62, 9/10, C12P 17/18, 19/32, C07D 498/18 // (C07D 498/18, 311:00, 273:00, 211:00)

(11) International Publication Number:

WO 00/20601

(43) International Publication Date:

13 April 2000 (13.04.00)

(21) International Application Number:

PCT/US99/22886

A3

(22) International Filing Date:

1 October 1999 (01.10.99)

(30) Priority Data:

60/102,748 2 October 1998 (02.10.98) US 60/123,810 11 March 1999 (11.03.99) US 60/139,650 17 June 1999 (17.06.99) US

(71) Applicant (for all designated States except US): KOSAN BIOSCIENCES, INC. [US/US]: 3832 Bay Center Drive, Hayward, CA 94545 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): REEVES, Christopher [US/US]; 4 East Altarinda Drive, Orinda, CA 94563 (US). CHU, Daniel [US/US]; 3767 Benton Street, Santa Clara, CA 95051 (US). KHOSLA, Chaitan [IN/US]; 740 Para Avenue, Palo Alto, CA 94306 (US). SANTI, Daniel [US/US]; 211 Belgrave Avenue, San Francisco, CA 94117 (US). WU, Kai [CN/US]; 900 Constitution Drive, Foster City, CA 94404 (US).

(74) Agents: FAVORITO, Carolyn et al.; Morrison & Foerster LLP, 2000 Pennsylvania Avenue, N.W., Washington, DC 20006–1888 (US).

(81) Designated States: AL, AM, AU, BA, BB, BG, BR, CA, CN, CR, CU, CZ, DM, EE, GD, GE, HR, HU, IL, IS, JP, KG, KP, KR, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, US, UZ, VN, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

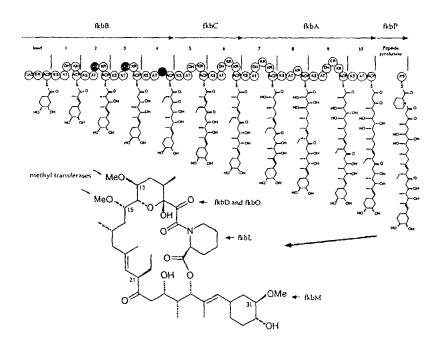
Published

With international search report.

(88) Date of publication of the international search report:

26 October 2000 (26.10.00)

(54) Title: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA CONSTRUCTS THEREFOR



(57) Abstract

Host cells comprising recombinant vectors encoding the FK-520 polyketide synthase and FK-520 modification enzymes can be used to produce the FK-520 polyketide. Recombinant DNA constructs comprising one or more FK-520 polyketide synthase domains, modules, open reading frames, and variants thereof can be used to produce recombinant polyketide synthases and a variety of different polyketides with application as pharmaceutical and veterinary products.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AΥ	Austria	FR	France	ւս	Luxembe	SN	Senegal
ΑU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Моласо	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
ВJ	Benin	ΙE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of Americ
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JР	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
Ct	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
Cf.	Cuba	ΚZ	Kazakstan	RO	Romania		
cz	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	Li	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

A. CLASSIFICATION OF SUBJECT MATTER | C12N15/52 | C12N15/54 | C12N15/62 | C12N9/10 | C12P17/18 | C12P19/32 | C07D498/18 | //(C07D498/18.311:00.273:00.211:00)

According to International Patent Classification (IPC) or to both national classification and (PC

B. FIELDS SEARCHED

Minimum gocumentation searched "classification system tollowed by classification symbols) IPC 7 C12N C12P C07D

Documentation searched other than minimum occumentation to the extent that such occuments are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where bractical, search terms used)

WPI Data, PAJ, MEDLINE, STRAND, BIOSIS, EMBASE, CHEM ABS Data

Category 3	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MOTAMEDI H ET AL.: "The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506" EUR. J. BIOCHEM., vol. 256, no. 3, 15 September 1998 (1998-09-15), pages 528-534, XP000906738 abstract figures 1,2 page 532, right-hand column, line 51 -page 533, left-hand column, line 18/	12

Further documents are listed in the continuation of box C.	Patent family members are listed in annex.		
"Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance. "E" earlier document but published on or after the international filling date. "L" document which may throw doubts on pnority claim(s) or which is cited to establish the publication date of another cration or other special reason (as specified). "O" document referring to an oral disclosure, use, exhibition or other means.	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention. "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone. "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.		
"P" document published prior to the international filing date but later than the priority date claimed	"&" document member of the same patent family		
Date of the actual completion of the international search	Date of mailing of the international search report		
27 July 2000	10/08/2000		
Name and mailing address of the ISA	Authorized officer		
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	van de Kamp, M		

INTERNATIONAL SEARCH KEPUKI

International Application No
PCT/US 99/22886

		PCT/US 99/22886			
	C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT Category Citation of document, with indication where appropriate, of the relevant passages (Relevant to claim No.)				
Sategory	Citation of document, with indication where appropriate, of the relevant passages	Relevant to claim No.			
X	STASSI D L ET AL.: "Ethyl-substituted erythromycin derivatives produced by directed metabolic engineering" PROC. NATL. ACAD. SCI. USA, vol. 95, no. 13, 23 June 1998 (1998-06-23). pages 7305-7309, XP002143632 abstract page 7308. left-hand column. line 4-right-hand column, line 17 page 7309, left-hand column, line 24-40	17			
X	REYNOLDS K A ET AL.: "Rapamycin, FK506, and ascomycin -related compounds" DRUGS PHARM. SCI., vol. 82, 1997, pages 497-520, XP000906777 figure 3 page 502, line 7-25; figure 7 page 509-513, paragraph IV	18			
X	DUMONT F J ET AL.: "The immunosuppressive and toxic effects of FK-506 are mechanistically related: pharmacology of a novel antagonist of FK-506 and rapamycin" JOURNAL OF EXPERIMENTAL MEDICINE, vol. 176, no. 3, 1 September 1992 (1992-09-01), pages 751-760, XP000906781 cited in the application abstract; figure 1	18			
X	KAWAI M ET AL.: "Structure-activity profiles of macrolactam immunosuppressant FK-506 analogues" FEBS LETTERS. vol. 316, no. 2, January 1993 (1993-01), pages 107-113, XP002143633 abstract scheme 1 table 1	18			
X	EP 0 323 042 A (FISONS PLC) 5 July 1989 (1989-07-05) example 13	18			
X	EP 0 356 399 A (SANDOZ AG ;SANDOZ AG (DE); SANDOZ AG (AT)) 28 February 1990 (1990-02-28) examples 2,3	18			
X	EP 0 463 690 A (MERCK & CO INC) 2 January 1992 (1992-01-02) example 3	18			
	-/				

C.(Continu	ation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication where appropriate, of the relevant passages	Relevant to claim No	
Y	MOTAMEDI H ET AL.: "Characterization of methyltransferase and hydroxylase genes involved in the biosynthesis of the immunosupressants FK506 and FK520" J. BACTERIOLOGY, vol. 178, no. 17, July 1996 (1996-07), pages 5243-5248, XP002137077 abstract page 5245, left-hand column, line 1-3 figure 4	1-11	
Y	MOTAMEDI H ET AL.: "Structural organization of a multi-unctional polyketide synthase involved in the biosynthesis of the macrolide immunosuppressant FK506" EUR. J. BIOCHEM., vol. 244, no. 1, 15 February 1997 (1997-02-15), pages 74-80, XP000906743 abstract page 79, left-hand column, line 26-35 page 75, left-hand column, line 31 -page 76, left-hand column, line 1	1-11	
A	CHEN T S ET AL.: "Microbial transformation of immunosuppressive compounds. II. Specific desmethylation of 13-methoxy group of FK 506 and FR 9500520 by Actinomycete sp. ATCC 53828" J. ANTIBIOT., vol. 45, no. 4, April 1992 (1992-04), pages 577-580, XP002143634 figure 1	18-20	
Α	SHAFIEE A ET AL.: "Enzymatic synthesis and immunosuppressive activity of novel desmethylated immunomycins (ascomycins)" J. ANTIBIOT., vol. 46, no. 9, September 1993 (1993-09), pages 1397-1405, XP002143635 abstract	18,20	
A	KHOSLA C: "Harnessing the biosynthetic potential of modular polyketide synthases" CHEMICAL REVIEWS, vol. 97, no. 7, 1997, pages 2577-2590, XP002130646 ISSN: 0009-2665		

MILLIAN DEPARTMENT ON

Information on patent family members

PCT/US 99/22886

Oatant do		Publication		Data da Camarilla		
Patent document cited in search report		date	Patent family member(s)		Publication date	
EP 0323	042 A	05-07-1989	AT	120466 T	15-04-1995	
			AU	630866 B	12-11-1992	
			AU	2822889 A	05-07-1989	
			CA	1339128 A	29-07-1997	
			CN	1033458 A	21-06-1989	
			DE	3853477 D	04-05-1995	
			DE	3853477 T	09-11-1995	
			DK	387889 A	08 -0 8-19 89	
			EΡ	03 4 6427 A	20-12-1989	
			ES	2071681 T	01-07-1995	
			FΙ	9 0 550 B	15-11-1993	
			FΙ	930597 A	11-02-1993	
			WO	8 905 304 A	15-06-1989	
			ΙE	66163 B	13-12-1995	
			IL	88629 A	12-04-1994	
			JP	2502463 T	09-08-1990	
			JP	2799208 B	17-09-1998	
			NO	893166 A	04-08-1989	
			NZ	227251 A	26-02-1990	
			PT	89203 A,B	29-12-1989	
			US	5376663 A	27-12-1994	
		ين جي مان سان بايان ويان وي من يوه سان مان بيان سان سان من دارد سان در	ZA	8809136 A	30-08-1989	
EP 0356	399 A	28-02-1990	AU	62 9 563 B	08-10-1992	
			AU	4024689 A	01-03-1990	
			DK	418789 A	27-02-1990	
			JP	2 167 287 A	27-06-1990	
			NZ	230418 A	25-10-1991	
			US	5011844 A	30-04-1991	
			ZA	8906524 A	24-04-1991	
EP 0463	690 A	02-01-1992	CA	2044846 A	26-12-1991	
			DE	69117221 D	28-03-1996	
			DE	69117221 T	12-09-1996	
			JP	4230288 A	19-08-1992	
			US	5190950 A	02-03-1993	
			US	5342935 A	30-08-1994	